In Class Assignment 8

November 6, 2024

1 In-Class Activity: Step-by-Step Machine Translation

Aaron Vo
CS 584
November 6, 2024

1.1 Objective

Translate the sentence ”The sun is shining” from English to French by coding each step in the translation process.

1.1.1 Step 1: Phrase Segmentation

Write a function to segment the sentence into predefined phrases. Task: Complete the function segment phrases to split ”The sun is shining” into the phrases [“The sun”, “is”, “shining”].

[6]:
```
import re
```

[15]:
```
def segment_phrases(sentence):
    words = sentence.split()
    phrases = [' '.join(words[:2])] + words[2:]
    return phrases
```

[16]:
```
sentence = "The sun is shining"
segments = segment_phrases(sentence)
print("Phrases: ", segments)
```

```
Phrases:  ['The sun', 'is', 'shining']
```

1.1.2 Step 2: Phrase Translation

Using a predefined dictionary, translate each phrase from English to French. Task: Complete the function translate phrases using the dictionary be- low.

[18]:
```
translation_dict = {
    "The sun": "Le soleil",
    "is": "est",
    "shining": "brille"
}
```

[20]:
```
def translate_phrases(phrases, translation_dict):
    translations = []
    for phrase in phrases:
        translations.append(translation_dict[phrase])
    return translations
```

[23]:
```
translations = translate_phrases(segments, translation_dict)
print("Translated Phrases: ", translations)
```

```
Translated Phrases:  ['Le soleil', 'est', 'brille']
```

1.1.3 Step 3: Reordering and Combining

Reorder and combine the translated phrases to form the final sentence in French. Task: Complete the function combine phrases to join the translated phrases.

[24]:
```
def combine_phrases(translations):
    return ' '.join(translations)
```

[25]:
```
final_sentence = combine_phrases(translations)
print("Final Translation: ", final_sentence)
```

```
Final Translation:  Le soleil est brille
```

1.1.4 Step 4: Probability Calculation

If time allows, assign probabilities to each phrase translation and calculate the overall probability of the sentence.

Task: Complete calculate probability to compute the likelihood of the translated sentence.

[26]:
```
phrase_probabilities = {
    "Le soleil": 0.9,
    "est": 0.95,
    "brille": 0.85
}

def calculate_probability(translations, phrase_probabilities):
    probability = 1.0
    for phrase in translations:
        probability *= phrase_probabilities[phrase]
    return probability
```

[27]:
```
print("Total Probability: ", calculate_probability(translations,␣
↪phrase_probabilities))
```

```
Total Probability:  0.72675
```