

```
from collections import Counter

with open("parsed_dependency.txt", "r", encoding="utf-8") as file:
    lines = file.readlines()

# store noun phrases
noun_phrases = []
np = []

print("DEBUG: Extracted words and UPOS tags:\n")

for line in lines:
    line = line.strip()

    # Skip empty lines and sentence headers
    if not line or line.startswith("Sentence:"):
        continue

    # Split line into word, lemma, and UPOS
    parts = line.split("\t")

    if len(parts) != 3:
        continue

    word, lemma, upos = parts

    print(f"Word: {word}, Lemma: {lemma}, UPOS: {upos}")

    if upos in ["NOUN", "ADJ"]:
        np.append(word)
    else:
        if np:
            if len(np) > 1:
                noun_phrases.append(" ".join(np))
            np = []

if np:
    noun_phrases.append(" ".join(np))

np_counts = Counter(noun_phrases)

# top 10 most frequent noun phrases
top_noun_phrases = np_counts.most_common(10)

# results
print("\nTop 10 Most Frequent Noun Phrases:")
for phrase, count in top_noun_phrases:
    print(f"{phrase}: {count}")
```