Grammar Design for Sentiment Analysis in NLP - Final Project

Seyedehmoojdeh Haghighat Hosseini
40171693

December 14, 2023

Parsing Complex Declarative Sentences

Special Verb Groups (SUBCAT)

These special verb groups in the grammar are really important for making detailed sentences. They decide how verbs work with other parts of the sentence like nouns, adverbs, and phrases that start with prepositions. Here are some main groups:

- ***_np_pp* Group:** This involves a verb followed by a noun and then a phrase starting with a preposition. It helps to make sentences where the action (verb) is connected to an object (noun) and then described more with a preposition. For example, "The director [VP] made [V] a film [NP] about love [PP]."
- ***_adjp* and *_np_adjp* Groups:** These groups include adjective phrases, right after a verb or inside a noun phrase. They let sentences give more details or qualities about someone or something. For example, "The movie [NP] was [VP] visually stunning [ADJP]."
- ***_s:that* Group:** This group helps to put a clause inside a verb phrase using words like "that." It's great for making complex sentences that give more information or context. For example, "The audience knew [VP] that the movie [S_THAT] would be a hit."

Extra Details (Relative Clauses)

Relative clauses add more information about nouns without starting a new sentence. They're key for making complex sentences in this grammar:

- **Standard Relative Clauses:** These clauses start with special pronouns like 'who,' 'which,' or 'that.' They add more details to simple sentences about the nouns. For example, "The movie [NP], which was directed by Nolan [RelClause], won several awards."
- **Relative Clauses without Direct Pronouns:** Sometimes, the clause starts right with a verb, especially when the special pronoun is understood but not said. This style is more common in everyday speaking.

Connecting Words (Coordinating Conjunctions)

The grammar's use of connecting words ('and,' 'but,' 'or') lets it join different parts of a sentence, each with its own idea or description. This is really useful for making sentences that express complex thoughts or show how different things are related.

Critique

While the grammar is comprehensive, several issues are notable:

1. **Complexity and Accessibility**: The detailed structure may be overwhelming and less accessible for non-experts.

2. **Ambiguity Handling**: The grammar does not explicitly address syntactic ambiguity.

3. **Not supporting the "zero" relative clauses**: In English, relative clauses can sometimes drop the relative pronoun when it is the object of the clause. For example, in the sentence "The movie we watched last night was very interesting," the relative pronoun "that" is omitted after "the movie." The full version would be "The movie that we watched last night was very interesting." Omitting relative pronouns can make sentences less complex and more natural, especially in spoken English. The inability of the grammar to parse these can lead to an incomplete understanding of common, everyday sentence structures and also senttiment

4. **Sentiment Analysis**: The inclusion of sentiment in syntactic structure is unconventional, typically considered a semantic aspect.

5. **Computational Efficiency**: The detailed and recursive nature might lead to inefficiencies in computational parsing.

6. **Lack of Semantic Context**: Focuses heavily on syntax with less emphasis on semantic relationships.

7. **Non-Standard Constructions**: Limited scope for parsing non-standard constructions, slang, or idiomatic expressions.

1 Handling of Sentiment-bearing Adjectives/Nouns with Conjunctions

Adjective and Noun Phrases

The grammar's structure includes rules for adjective phrases (ADJP) and noun phrases (NP) where sentiment ('positive', 'negative', 'neutral') is a key attribute. This allows for capturing the sentiment of individual adjectives or nouns.

1.0.1 Use of Coordinating Conjunctions (CC)

The grammar incorporates rules for coordinating conjunctions (CC) like 'and', 'or', and 'but'. This is essential for linking multiple sentiment-bearing adjectives or nouns. For example, in "a dull and scary movie," the grammar would parse "dull" and "scary" as separate adjective phrases joined by 'and'.

1.0.2 Sentiment Analysis in Conjunctions

The grammar seems to allow for sentiment analysis within phrases connected by conjunctions. For example, in "It was a mess and a hazard," both "mess" and "hazard" can be identified as negative sentiments linked by 'and'.

1.0.3 Handling Negation

The grammar includes rules for negation ('not', "n't") which can invert the sentiment of an adjective or a noun. This is crucial for sentences where negation changes the overall sentiment, such as "It was not a mess or a hazard". It is done by Verb negation an Helper negation part of the grammar.

1.0.4 Critique of the Grammar Design

- **Ambiguity in Sentiment Analysis**: The grammar does not explicitly address how it would handle ambiguity in sentiment, particularly when adjectives or nouns have context-dependent meanings.
- **Over-reliance on Syntax for Sentiment**: Relying primarily on syntactic structure for sentiment inference may not always capture language nuances where context is key.
- **Computational Efficiency**: The detailed and recursive nature of this grammar might be inefficient for computational parsing, potentially leading to increased processing time and resource usage.
- **Lack of Semantic Context**: There's potential inadequacy in integrating broader semantic context, which is crucial for accurate sentiment analysis.

Handling of Sentiment-bearing Sentences

Sentence Structure

The grammar defines a sentence (`S`) as a combination of noun phrases (`NP`) and verb phrases (`VP`), with optional modal verbs and negation. Each `NP` and `VP` can carry a sentiment value (*positive*, *negative*, *neutral*).

Handling of Coordinating Conjunctions (`CC`)

The grammar uses coordinating conjunctions (*and*, *or*, *but*) to link multiple clauses or phrases, each potentially carrying its sentiment. For example, in "It was too long but entertaining," the grammar can parse "too long" and "entertaining" as separate phrases joined by *but*.

Negation

The grammar includes rules for negation (*not*, *n't*), which is crucial in sentences where negation alters the sentiment. The structure ensures that the sentiment is correctly interpreted in the context of negation.

Critique of the Grammar Design

- **Complexity**: While the grammar is comprehensive in handling conjunctions and sentiments, its complexity could be a challenge, especially with sentences having multiple layers of conjunctions and negations.
- **Ambiguity and Context Sensitivity**: The grammar might struggle with ambiguous cases where the sentiment of a phrase depends heavily on the context, which is not always syntactically evident.
- **Computational Efficiency**: The recursive nature of this grammar, particularly in parsing complex sentences with multiple conjunctions and negations, might be computationally intensive, leading to longer processing times.
- **Semantic Depth**: The grammar primarily focuses on syntactic structures and may not fully capture the semantic nuances essential for accurate sentiment analysis.
- **Handling of Complex Negations**: The grammar might struggle with sentences where negation is used more complexly, such as double negatives or indirect negation.

Comparison of SSAP and Grammar-Based Analysis

Sentence	Grammar-based	SSAP	SSAP Sentiment Score
It was too long but entertaining	Positive	Positive	2
I saw a dull and scary movie, it was a mess and a hazard	Negative	Negative	-6
This may not have the dramatic gut-wrenching impact of other holocaust films, but it is a compelling story	Negative	Neutral	0
A perfect example of rancid, well-intentioned, but shamelessly manipulative movie making	Positive	Positive	3
He passed the test which was quite a surprise.	Neutral	Neutral	0
Manipulative movie making	Negative	Neutral	0

Table 1: Performance Comparison of SSAP and grammar-based sentiment analysis

Error Analysis of SSAP and Project 4

It was too long but entertaining

SSAP: Accurately captured the positive sentiment.
Project 4: Also accurately captured the positive sentiment, but with a lower sentiment score, indicating less confidence or different weighting.

I saw a dull and scary movie, it was a mess and a hazard

SSAP: Correctly identified as negative with a strong sentiment score.
Project 4: Also identified as negative but with a less negative score, suggesting a less intense sentiment perception.

This may not have the dramatic gut-wrenching impact of other holocaust films, but it is a compelling story

SSAP: Predicted as neutral, possibly due to conflicting sentiments in the sentence.
Project 4: Predicted as positive, indicating a possible focus on the latter part of the sentence which has a positive sentiment.

A perfect example of rancid, well-intentioned, but shamelessly manipulative movie making

SSAP: Identified as positive, potentially focusing on terms like "perfect" and "well-intentioned".
Project 4: Identified as negative, which could indicate a focus on the negative terms like "rancid" and "manipulative".

Manipulative movie making

SSAP: Predicted as neutral, possibly due to the lack of clear sentiment words.
Project 4: Correctly identified as negative, likely recognizing "manipulative" as a negative term.

Stance Assignment

1.1 Approach for Stance Assignment in Movie Reviews

The grammar designed for parsing sentences in movie reviews is structured to identify the sentiment towards named entities, particularly focusing on various syntactic elements. This section breaks down the key components of the grammar and its approach to stance assignment:

- **Sentence Structure (S):** The grammar defines a sentence (S) to include noun phrases (NP) and verb phrases (VP), each carrying a sentiment attribute ('SENT=?s'). This structure is essential for determining the overall sentiment of the sentence.
- **Noun Phrases (NP) and Proper Nouns (NNP):** Named entities, like movie titles or actor names, are often captured as Noun Phrases or Proper Nouns. These are key to identifying the entity towards which the sentiment is directed.
- **Use of Adjective Phrases (ADJP) in NPs:** Adjective Phrases within Noun Phrases capture descriptive elements about the named entities, contributing to the overall stance towards the entity.
- **Verb Phrases (VP) and Sentiment:** Verb Phrases are crucial for understanding the actions or states associated with the named entities, with their sentiment aiding in stance determination.
- **Handling of Coordinating Conjunctions (CC):** The grammar's treatment of conjunctions like 'and', 'or', 'but' allows for complex sentences with multiple sentiments or entities.
- **Relative Clauses and Complex Sentences:** Inclusion of relative clauses enables the grammar to handle complex sentences that provide additional information about the named entity, influencing sentiment.
- **Negation and Modal Verbs:** Rules for negation and modal verbs are critical in accurately capturing the sentiment, especially when it is inverted.
- **Adverbial Phrases (ADVP):** Used to modify the sentiment expressed, amplifying or altering the stance towards an entity.
- **Integration of Lexicon for Sentiment Analysis:** Assumes the integration with a sentiment analyzer, essential for quantifying the sentiments expressed in the reviews.

In essence, the grammar analyzes movie reviews at a detailed syntactic level, focusing on noun phrases for identifying entities and using other grammatical elements to determine the sentiment or stance towards these entities. The integration of a sentiment lexicon is crucial, and the effectiveness of this approach heavily depends on its comprehensiveness and accuracy.