USING RHETORICAL STRUCTURE IN SENTIMENT ANALYSIS OF TEXT

ALEXANDER HOGENBOOM

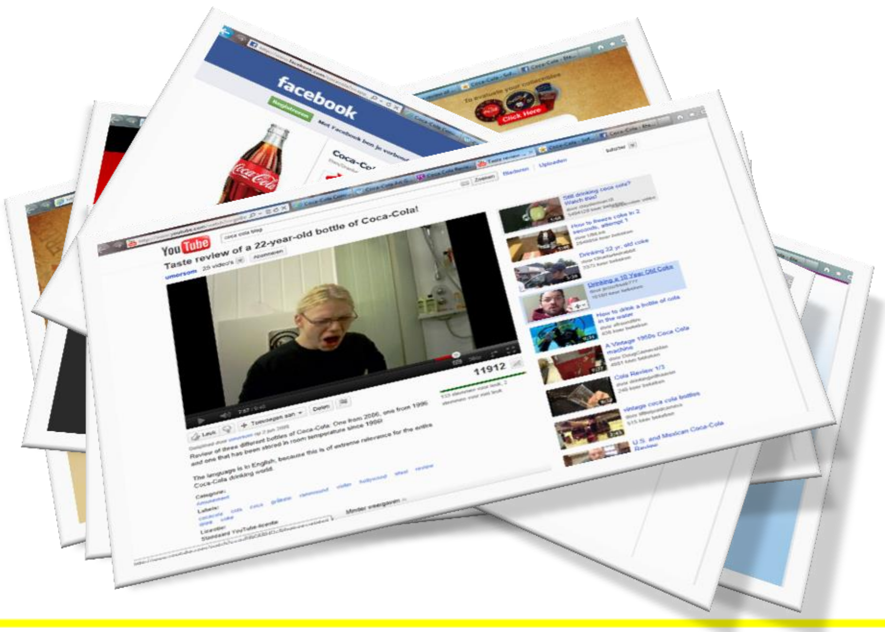
Full paper entitled "*Using Rhetorical Structure in Sentiment Analysis*", authored by Alexander Hogenboom, Flavius Frasincar, Franciska de Jong, and Uzay Kaymak, will appear in *Communications of the ACM*.

INTRODUCTION (1)

- Semantic information systems for decision support can benefit from a sense of people’s sentiment with respect to events, products, brands, etc.
- The Web offers an overwhelming amount of textual data, containing traces of sentiment

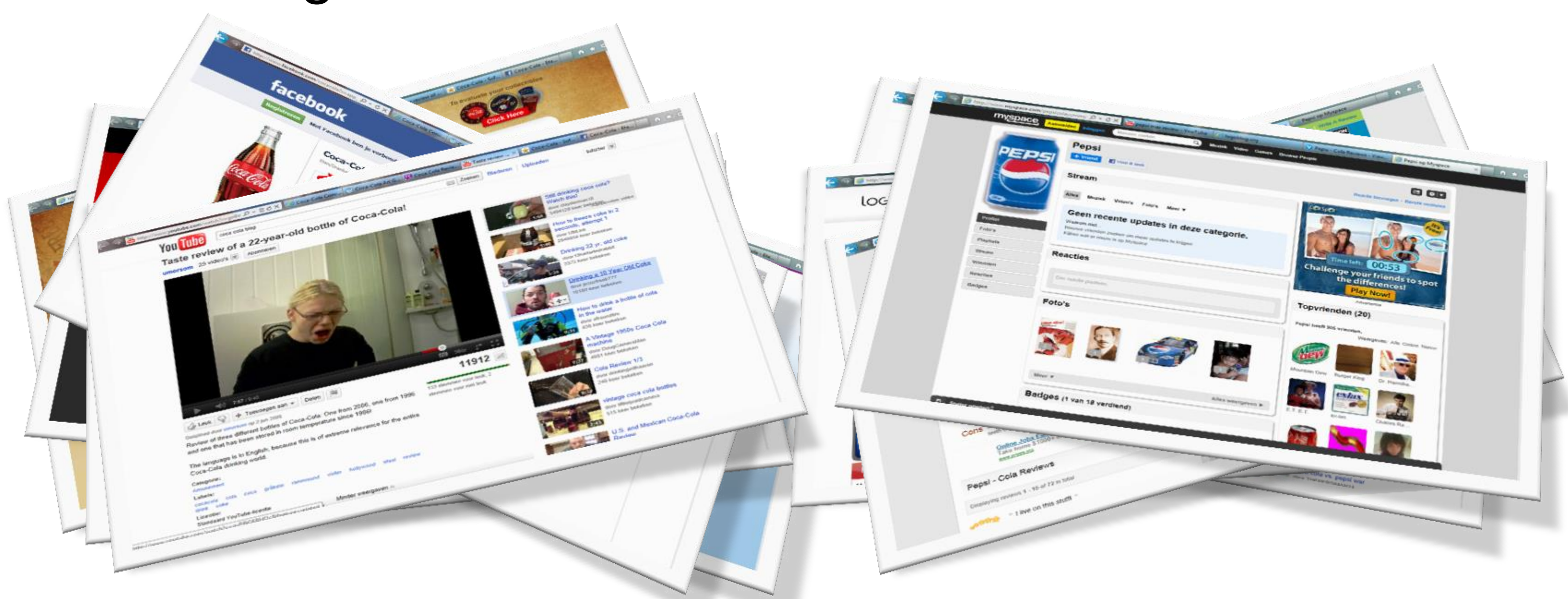
INTRODUCTION (1)

- Semantic information systems for decision support can benefit from a sense of people's sentiment with respect to events, products, brands, etc.
- The Web offers an overwhelming amount of textual data, containing traces of sentiment



INTRODUCTION (1)

- Semantic information systems for decision support can benefit from a sense of people’s sentiment with respect to events, products, brands, etc.
- The Web offers an overwhelming amount of textual data, containing traces of sentiment



INTRODUCTION (2)

- An intuitive approach to sentiment analysis involves scanning a text for cues signaling its polarity

INTRODUCTION (2)

- An intuitive approach to sentiment analysis involves scanning a text for cues signaling its polarity
- Let us consider the following *negative* review:
  - Example: *Although Brad Pitt's well-deserved fall off a cliff was quite entertaining, this movie was terrible!*

INTRODUCTION (2)

- An intuitive approach to sentiment analysis involves scanning a text for cues signaling its polarity
- Let us consider the following *negative* review:
  - Example: *Although Brad Pitt's well-deserved fall off a cliff was quite entertaining, this movie was terrible!*
- How can the structure of natural language text be exploited when determining the text's polarity?

INTRODUCTION (2)

- An intuitive approach to sentiment analysis involves scanning a text for cues signaling its polarity
- Let us consider the following *negative* review:
  - Example: *Although Brad Pitt's well-deserved fall off a cliff was quite entertaining, this movie was terrible!*
- How can the structure of natural language text be exploited when determining the text's polarity?

SENTIMENT ANALYSIS

- Sentiment analysis is typically focused on determining the polarity of natural language text
- Applications in summarizing reviews, determining a general mood (consumer confidence, politics)

SENTIMENT ANALYSIS

- Sentiment analysis is typically focused on determining the polarity of natural language text
- Applications in summarizing reviews, determining a general mood (consumer confidence, politics)
- State-of-the-art approaches classify the polarity of natural language text by analyzing vector representations using, e.g., machine learning techniques
- Alternative approaches are lexicon-based, which renders them robust across domains and texts and enables linguistic analysis at a deeper level

STRUCTURE-GUIDED CLASSIFICATION (1)

- Early approaches involve accounting for segments' positions in a text or their semantic cohesion
- Recent work exploits discursive relations by applying the Rhetorical Structure Theory (RST)

STRUCTURE-GUIDED CLASSIFICATION (1)

- Early approaches involve accounting for segments' positions in a text or their semantic cohesion
- Recent work exploits discursive relations by applying the Rhetorical Structure Theory (RST)
- RST can be used to split a text into a hierarchical structure of rhetorically related segments
- Nucleus segments form the core of a text, whereas satellites support the nuclei
- Many types of relations between segments exist, e.g., background, elaboration, explanation, contrast, etc.

STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:

STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:

While always *complaining* that	he *hates* this type of movies,	John *bitterly* confessed that	he *enjoyed* this movie.

STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:

While always *complaining* that | he *hates* this type of movies, | John *bitterly* confessed that | he *enjoyed* this movie.

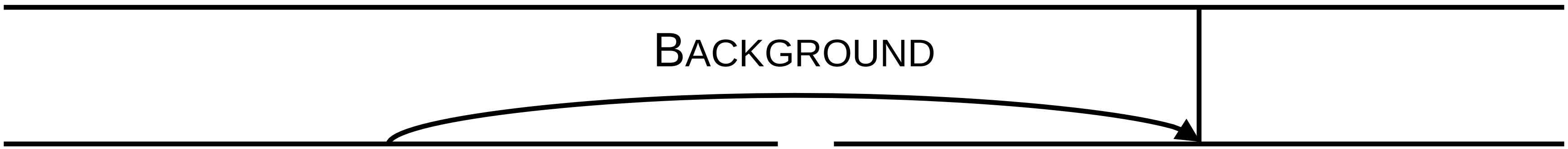
STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:

While always *complaining* that | he *hates* this type of movies, | John *bitterly* confessed that | he *enjoyed* this movie.

STRUCTURE-GUIDED CLASSIFICATION (2)

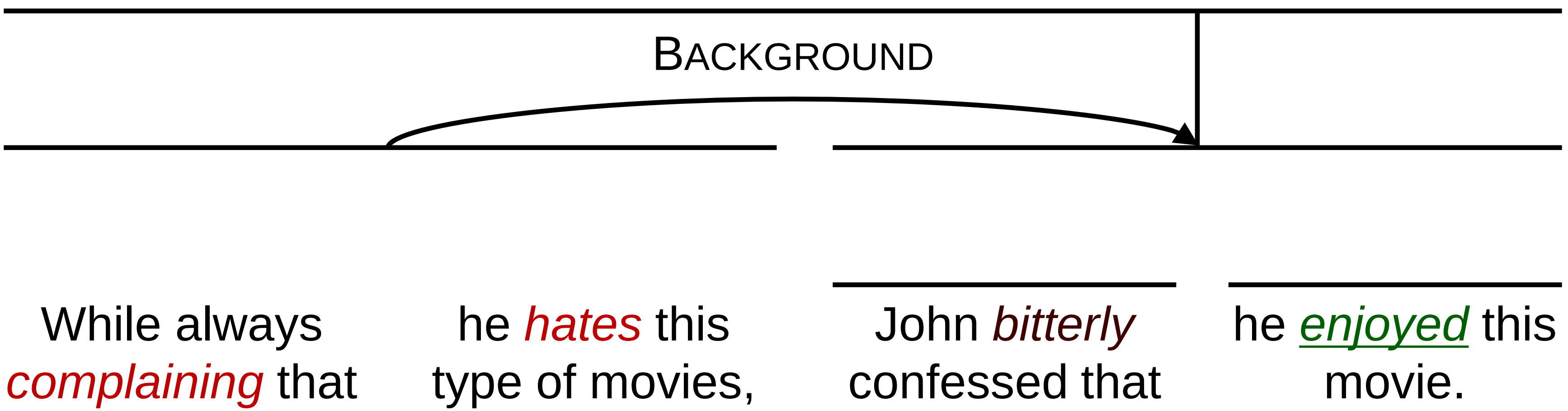
- Rhetorical structure of an example sentence:



While always *complaining* that	he *hates* this type of movies,	John *bitterly* confessed that	he *enjoyed* this movie.

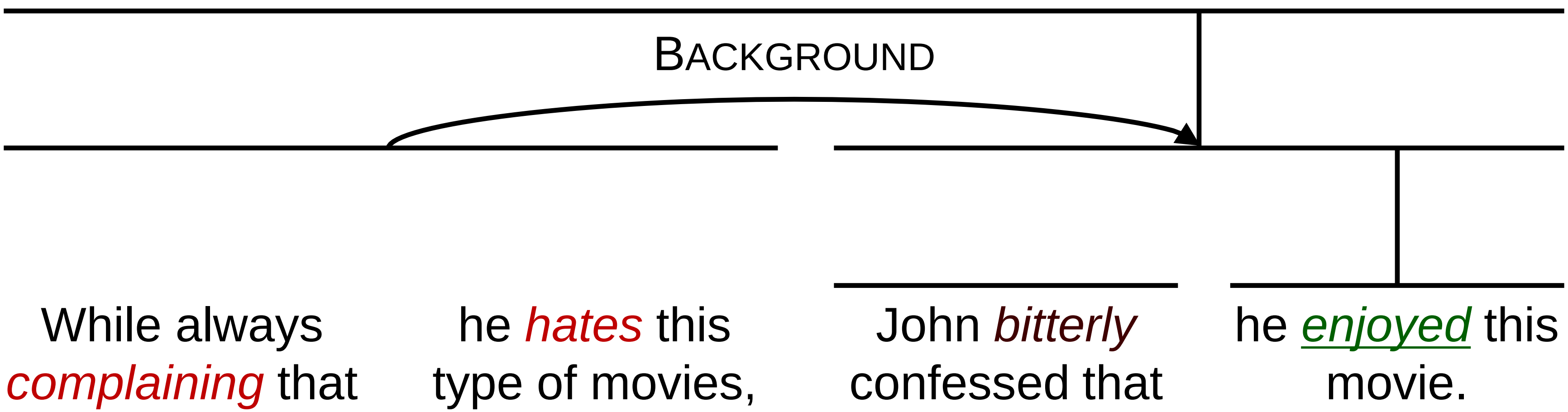
STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:



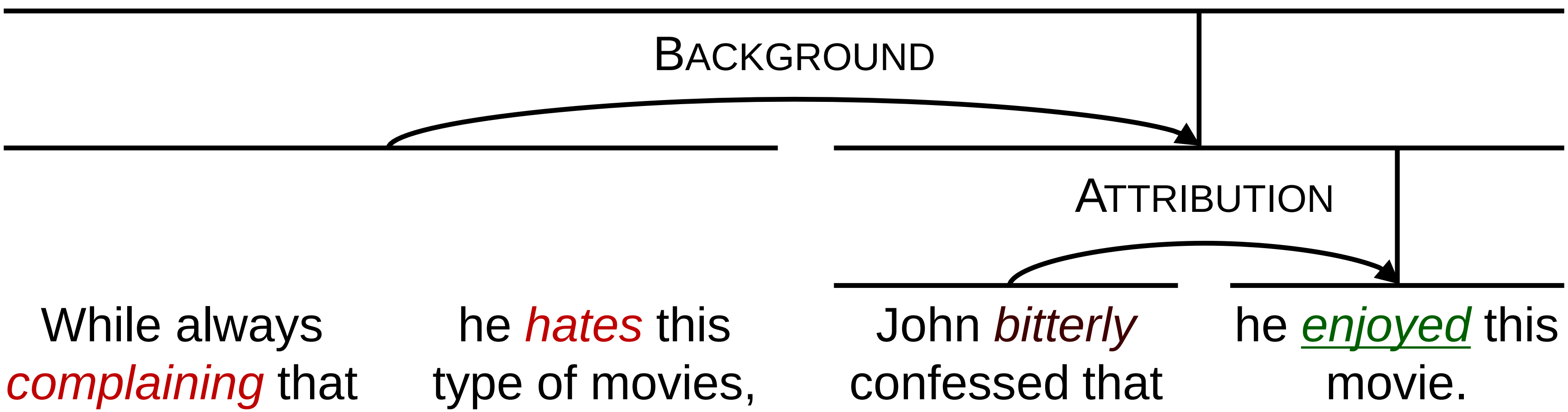
STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:



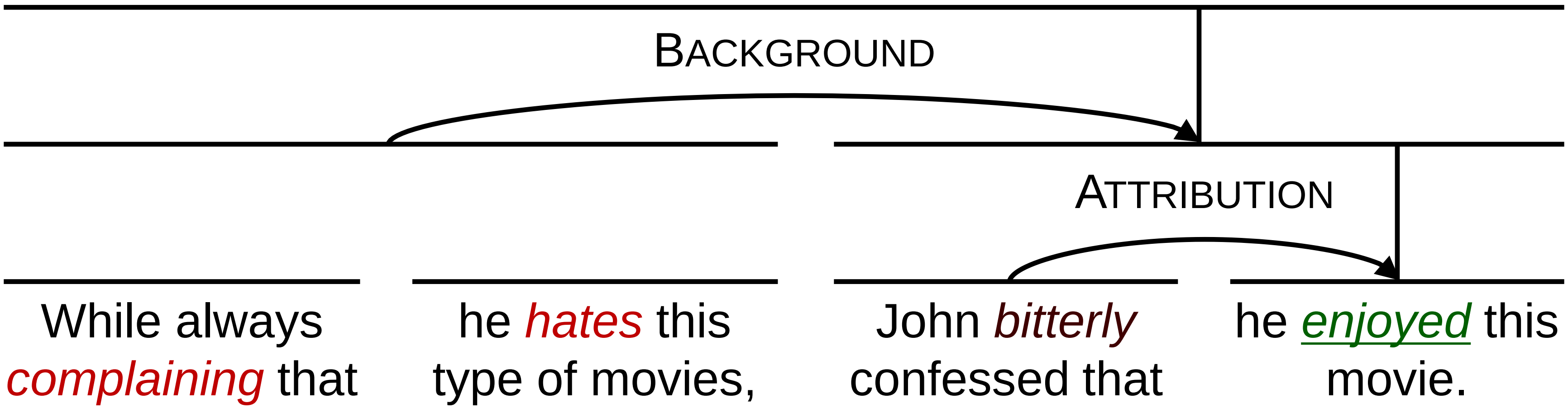
STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:



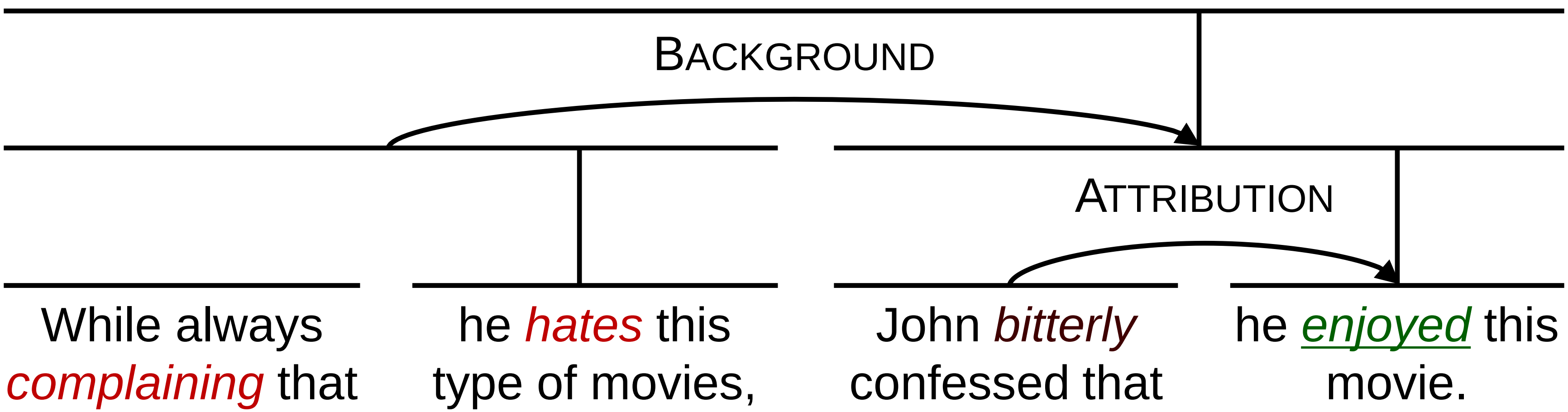
STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:



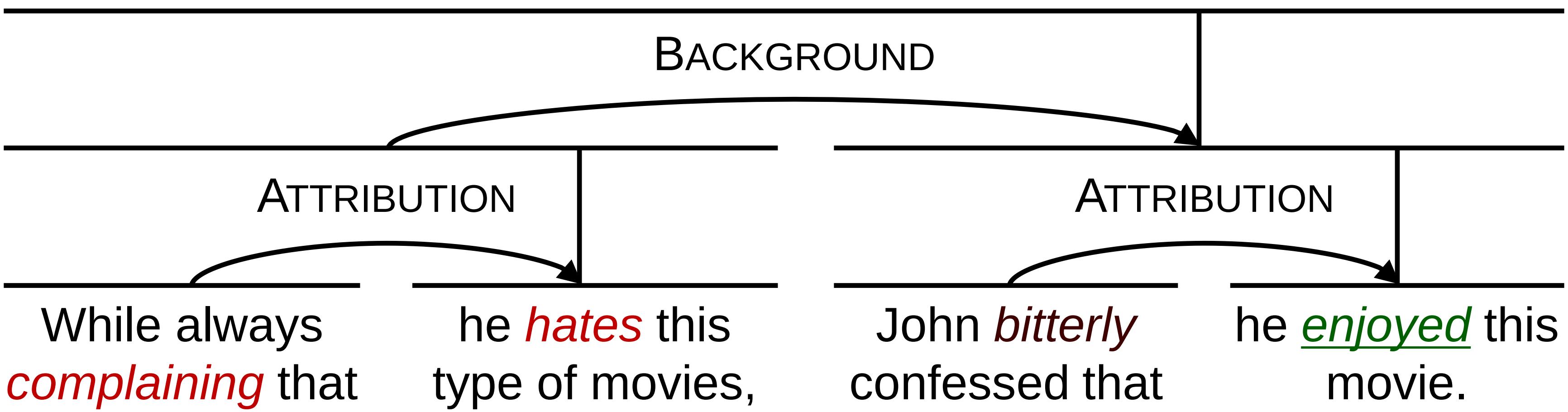
STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:



STRUCTURE-GUIDED CLASSIFICATION (2)

- Rhetorical structure of an example sentence:



STRUCTURE-GUIDED CLASSIFICATION (3)

- Existing work differentiates between important and less important segments w.r.t. the overall sentiment
- Previously proposed methods assign different weights to top-level nuclei and satellites in sentences

STRUCTURE-GUIDED CLASSIFICATION (3)

- Existing work differentiates between important and less important segments w.r.t. the overall sentiment
- Previously proposed methods assign different weights to top-level nuclei and satellites in sentences
- We propose to:

STRUCTURE-GUIDED CLASSIFICATION (3)

- Existing work differentiates between important and less important segments w.r.t. the overall sentiment
- Previously proposed methods assign different weights to top-level nuclei and satellites in sentences
- We propose to:
  - Differentiate among rhetorical roles, i.e., RST relation types

STRUCTURE-GUIDED CLASSIFICATION (3)

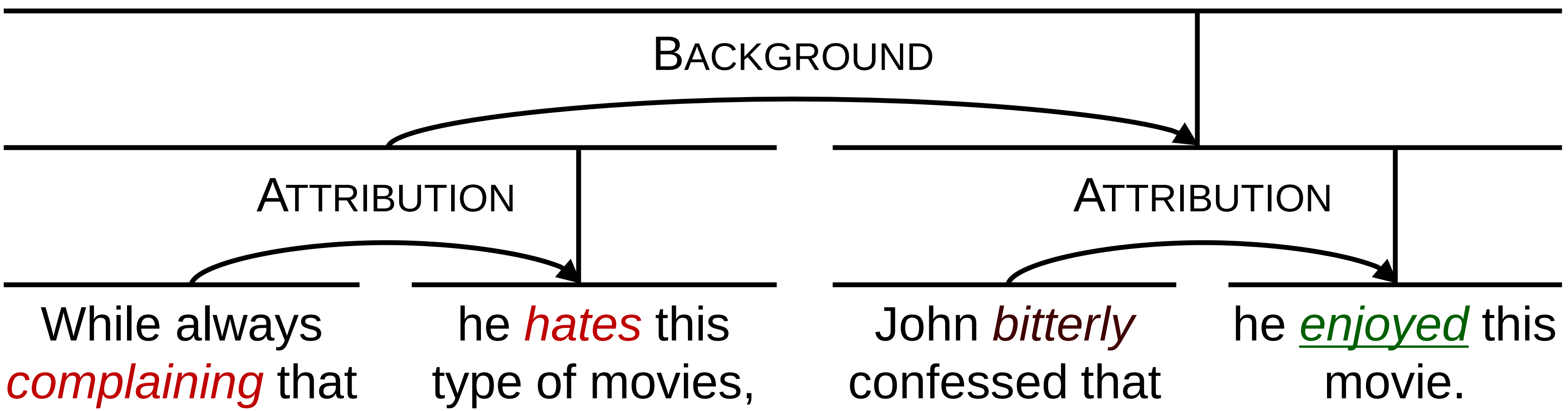
- Existing work differentiates between important and less important segments w.r.t. the overall sentiment
- Previously proposed methods assign different weights to top-level nuclei and satellites in sentences
- We propose to:
  - Differentiate among rhetorical roles, i.e., RST relation types
  - Account for the full rhetorical structure

STRUCTURE-GUIDED CLASSIFICATION (3)

- Existing work differentiates between important and less important segments w.r.t. the overall sentiment
- Previously proposed methods assign different weights to top-level nuclei and satellites in sentences
- We propose to:
  - Differentiate among rhetorical roles, i.e., RST relation types
  - Account for the full rhetorical structure
  - Guide polarity classification by sentence-, paragraph-, and document-level RST analysis

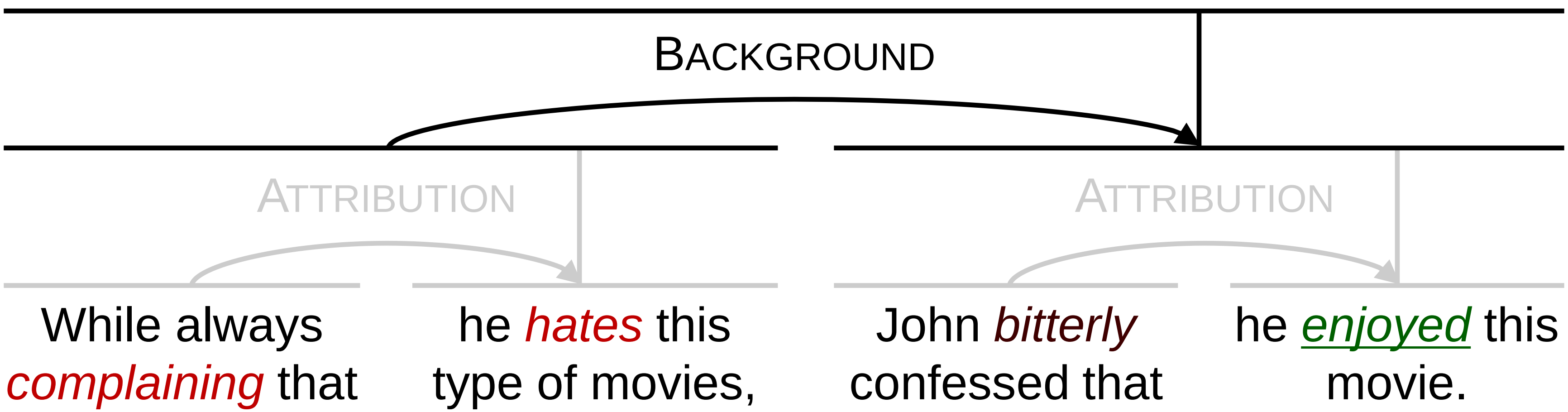
STRUCTURE-GUIDED CLASSIFICATION (4)

- Top-level RST-guided sentiment analysis:



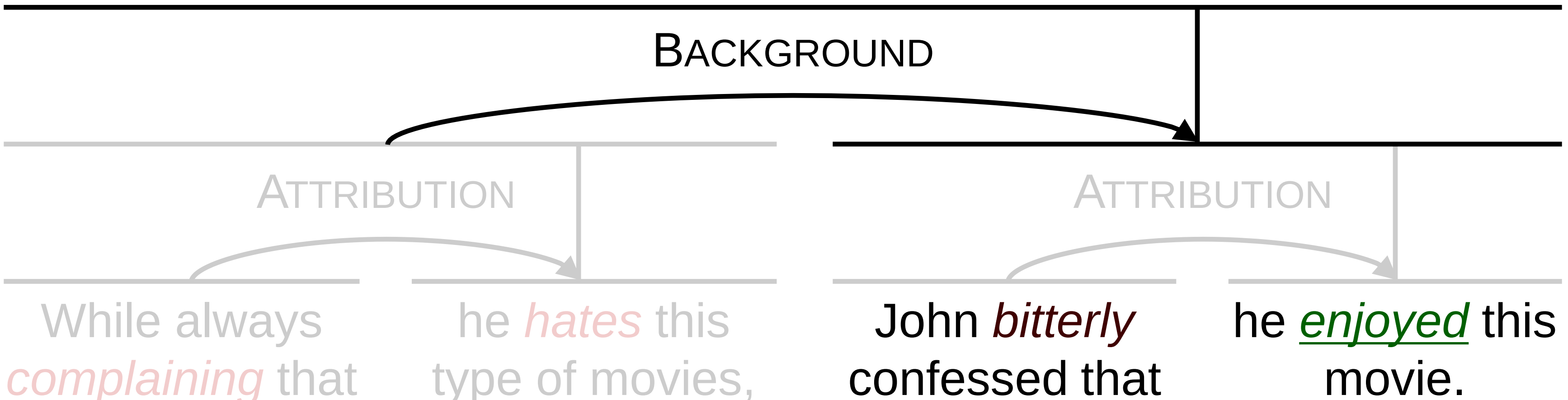
STRUCTURE-GUIDED CLASSIFICATION (4)

- Top-level RST-guided sentiment analysis:



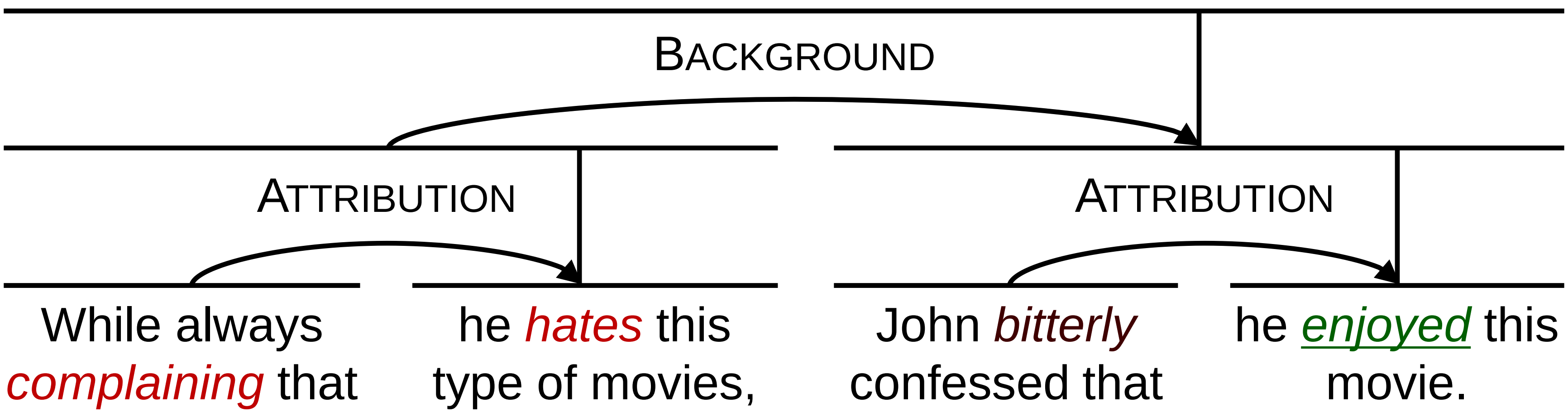
STRUCTURE-GUIDED CLASSIFICATION (4)

- Top-level RST-guided sentiment analysis:



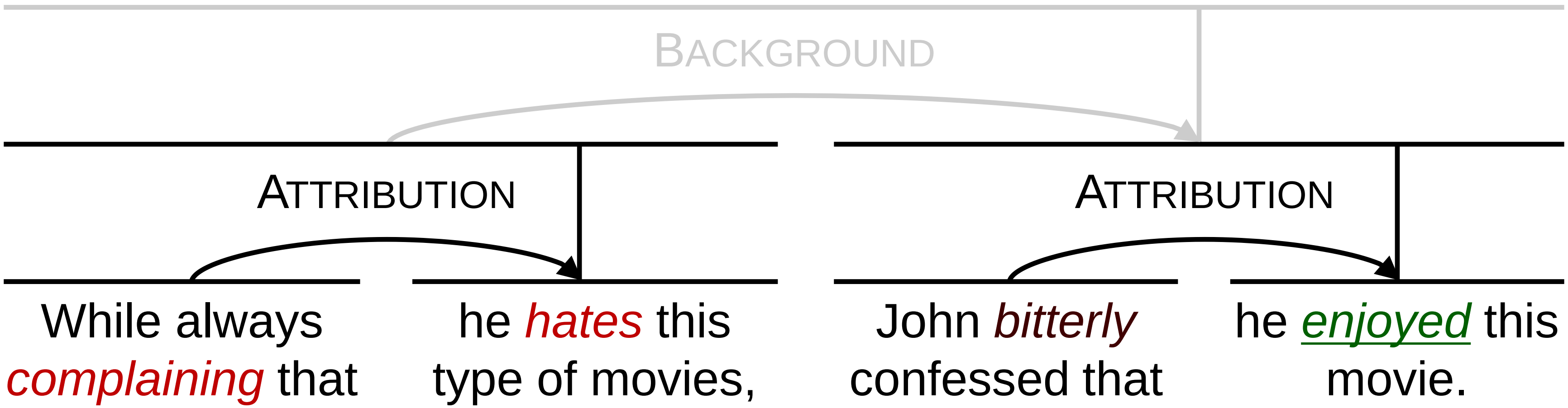
STRUCTURE-GUIDED CLASSIFICATION (5)

- Leaf-level RST-guided sentiment analysis:



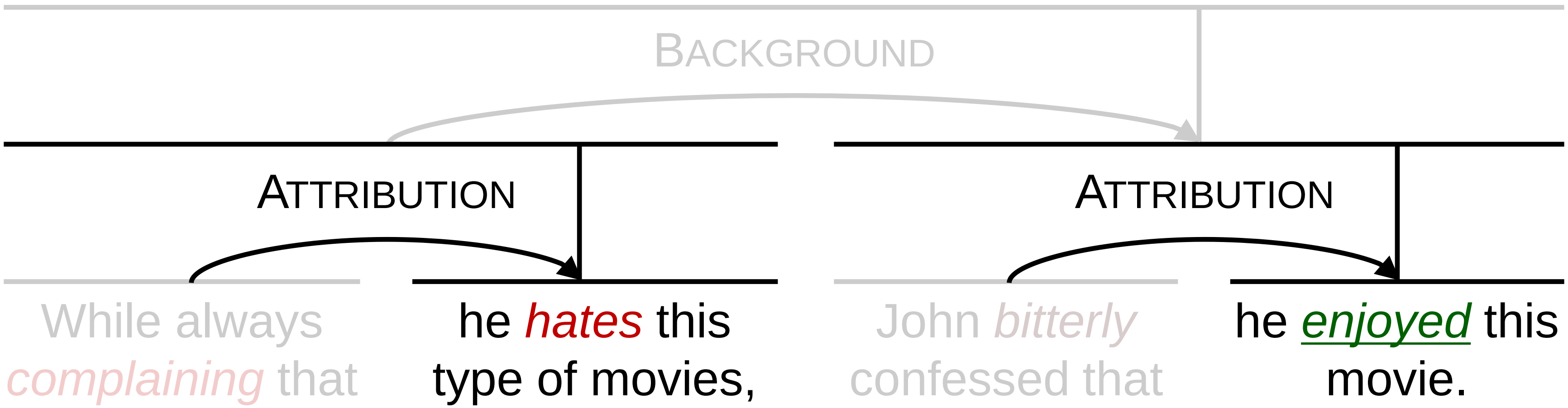
STRUCTURE-GUIDED CLASSIFICATION (5)

- Leaf-level RST-guided sentiment analysis:



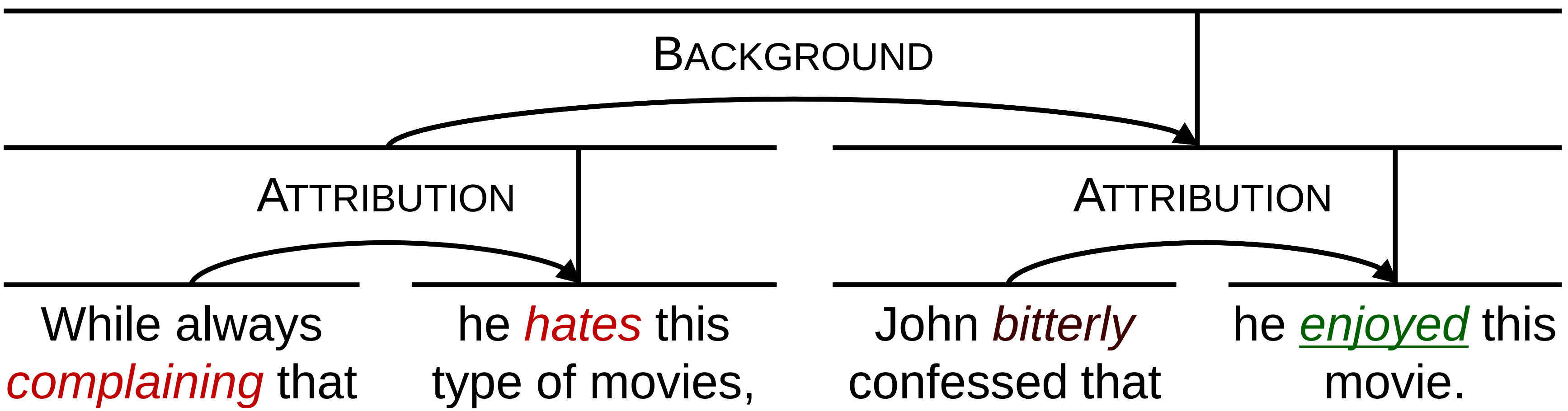
STRUCTURE-GUIDED CLASSIFICATION (5)

- Leaf-level RST-guided sentiment analysis:



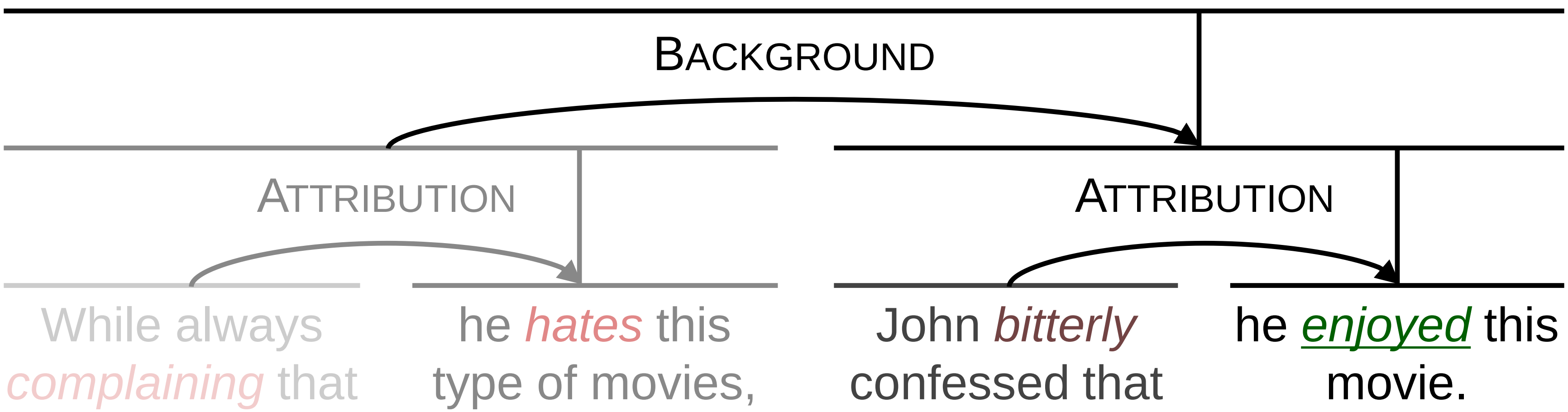
STRUCTURE-GUIDED CLASSIFICATION (6)

- Hierarchical RST-guided sentiment analysis:



STRUCTURE-GUIDED CLASSIFICATION (6)

- Hierarchical RST-guided sentiment analysis:



FRAMEWORK (1)

- Lexicon-based document-level polarity classification
- Based on its lemma, Part-of-Speech (POS), and disambiguated word sense (Lesk-based), each individual word is scored in the range [-1,1]

FRAMEWORK (1)

- Lexicon-based document-level polarity classification
- Based on its lemma, Part-of-Speech (POS), and disambiguated word sense (Lesk-based), each individual word is scored in the range [-1,1]
- Word scores are aggregated and corrected for a bias towards positivity in order to classify text as positive (corrected score ≥ 0) or negative (corrected score < 0)
- Discourse parsing is applied in order to determine appropriate weights for word scores in this process

FRAMEWORK (2)

- Simple discourse parsing: weights proportional to position of words in the full text

FRAMEWORK (2)

- Simple discourse parsing: weights proportional to position of words in the full text
- RST-based discourse parsing:

FRAMEWORK (2)

- Simple discourse parsing: weights proportional to position of words in the full text
- RST-based discourse parsing:
  - Unit of analysis: sentence (S), paragraph (P), or document (D)

FRAMEWORK (2)

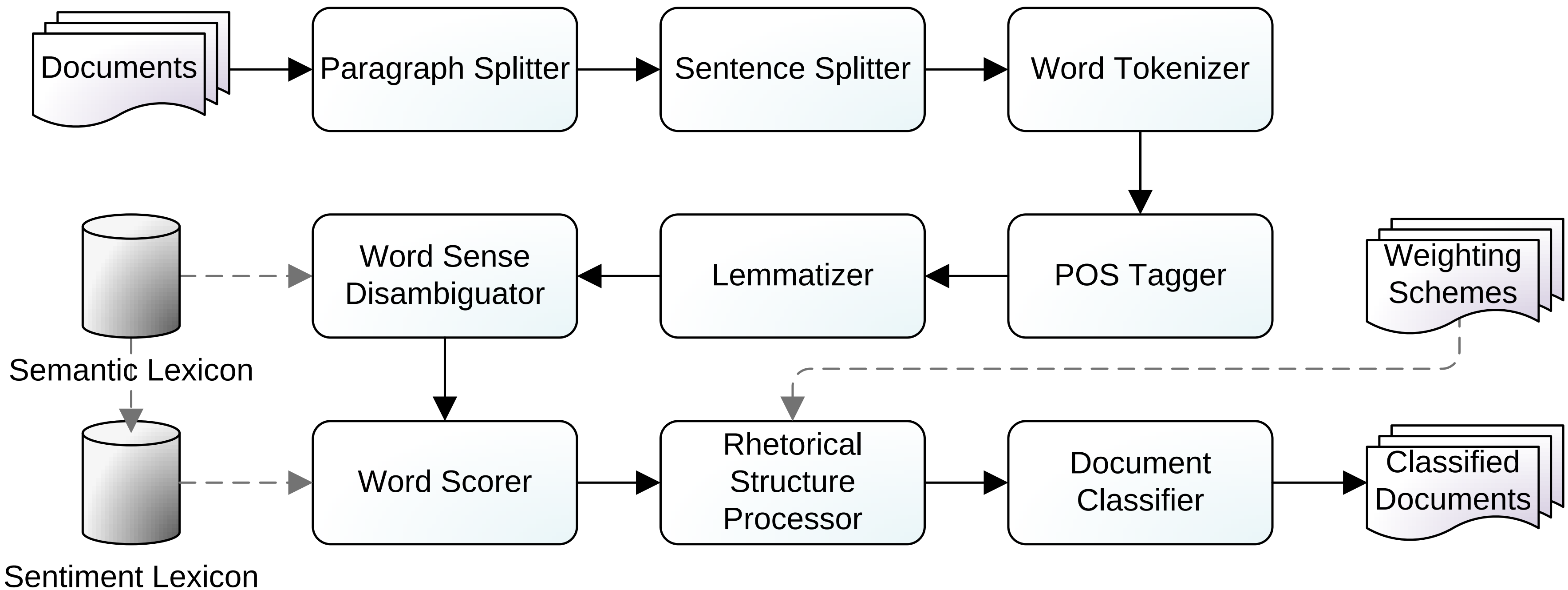
- Simple discourse parsing: weights proportional to position of words in the full text
- RST-based discourse parsing:
  - Unit of analysis: sentence (S), paragraph (P), or document (D)
  - Rhetorical parsing: top-level (T), leaf-level (L), or hierarchical (H)

FRAMEWORK (2)

- Simple discourse parsing: weights proportional to position of words in the full text
- RST-based discourse parsing:
  - Unit of analysis: sentence (S), paragraph (P), or document (D)
  - Rhetorical parsing: top-level (T), leaf-level (L), or hierarchical (H)
  - Weighting schemes:
    - I: nucleus weights of 1, satellite weights of 0
    - II: nucleus weights of 1.5, satellite weights of 0.5
    - X: optimized weights, differentiated by satellite type
    - F: optimized weights, differentiated by nucleus and satellite type

FRAMEWORK (3)

- Polarity classification framework:



EVALUATION (1)

- Implementation in Java, Stanford tokenizer, OpenNLP POS tagger, Java WordNet Library (JWNL) API for lemmatization, SentiWordNet 3.0 sentiment lexicon
- RST parsers:
  - Sentence-level PArsing of DiscoursE (SPADE, for sentences)
  - HIgh-Level Discourse Analyzer (HILDA, for sentences, paragraphs, and documents)

EVALUATION (1)

- Implementation in Java, Stanford tokenizer, OpenNLP POS tagger, Java WordNet Library (JWNL) API for lemmatization, SentiWordNet 3.0 sentiment lexicon
- RST parsers:
  - Sentence-level PArsing of DiscoursE (SPADE, for sentences)
  - HIgh-Level Discourse Analyzer (HILDA, for sentences, paragraphs, and documents)
- Optimization of weights for weighting schemes X and F by means of a Particle Swarm Optimization approach
- For hierarchical RST-based sentiment analysis, a diminishing factor is optimized as well

EVALUATION (2)

- Performance evaluation by 10-fold cross-validation on a corpus of 1,000 positive and 1,000 negative manually classified English movie reviews

EVALUATION (2)

- Performance evaluation by 10-fold cross-validation on a corpus of 1,000 positive and 1,000 negative manually classified English movie reviews
- Baselines:
  - All words assigned a weight of 1 (no RST-guided analysis)
  - Simple discourse parsing (position-based analysis)

EVALUATION (2)

- Performance evaluation by 10-fold cross-validation on a corpus of 1,000 positive and 1,000 negative manually classified English movie reviews
- Baselines:
  - All words assigned a weight of 1 (no RST-guided analysis)
  - Simple discourse parsing (position-based analysis)
- RST-guided alternatives are all 48 combinations of:
  - Unit of analysis S (for both SPADE and HILDA), P, and D
  - Parsing methods T, L, and H
  - Weighting schemes I, II, X, and F

EVALUATION (3)

- Performance of baselines and best methods for each level of analysis:

	Positive			Negative			Overall	
Method	Prec.	Rec.	F1	Prec.	Rec.	F1	Acc.	F1
Baseline	.632	.689	.659	.658	.599	.627	.644	.643
Position	.637	.713	.673	.674	.593	.631	.653	.652
SPADE.S H F	.710	.738	.724	.727	.699	.713	.719	.718
HILDA.S L F	.705	.732	.718	.721	.693	.707	.713	.712
HILDA.P H F	.713	.692	.703	.701	.722	.711	.707	.707
HILDA.D L F	.701	.727	.714	.717	.690	.703	.709	.708

EVALUATION (4)

EVALUATION (4)

- Sentence-level RST trees are most useful in document-level polarity classification

EVALUATION (4)

- Sentence-level RST trees are most useful in document-level polarity classification
- Deeper analysis of sentential rhetorical structure helps significantly improve polarity classification performance

EVALUATION (4)

- Sentence-level RST trees are most useful in document-level polarity classification
- Deeper analysis of sentential rhetorical structure helps significantly improve polarity classification performance
- Weighting schemes X and F significantly outperform the baselines and the other weighting schemes

EVALUATION (4)

- Sentence-level RST trees are most useful in document-level polarity classification
- Deeper analysis of sentential rhetorical structure helps significantly improve polarity classification performance
- Weighting schemes X and F significantly outperform the baselines and the other weighting schemes
- Nuclei are comparably important, but some satellites are important as well (e.g., contrasting or elaborating ones)

EVALUATION (4)

- Sentence-level RST trees are most useful in document-level polarity classification
- Deeper analysis of sentential rhetorical structure helps significantly improve polarity classification performance
- Weighting schemes X and F significantly outperform the baselines and the other weighting schemes
- Nuclei are comparably important, but some satellites are important as well (e.g., contrasting or elaborating ones)
- Optimized diminishing factors result in the first 15 to 30 levels of an RST tree to be accounted for in the analysis

EVALUATION (4)

- Sentence-level RST trees are most useful in document-level polarity classification
- Deeper analysis of sentential rhetorical structure helps significantly improve polarity classification performance
- Weighting schemes X and F significantly outperform the baselines and the other weighting schemes
- Nuclei are comparably important, but some satellites are important as well (e.g., contrasting or elaborating ones)
- Optimized diminishing factors result in the first 15 to 30 levels of an RST tree to be accounted for in the analysis
- The best approaches enable a focused analysis

EVALUATION (5)

We 're back in blade runner territory with this one , *conceptual* artist robert longo 's *vision* of a william *gibson-inspired* future where information *is* the commodity to *kill* for . Front and center *is* johnny (keanu reeves) , a " cyber-courier " who *smuggles* data via a " wet-wired " implant . He 's ready to quit the biz and *get* a portion of his long-term *memory* restored , but , first , he has to finish one *last* , *dangerous* job .

The pressing *problem* in johnny mnemonic *is* that keanu reeves seems to have *forgotten* how to play an action hero since his stint on speed . He 's walking wood in a forest of stiffs that includes henry rollins , ice-t , and dina meyer . (dolph lundgren 's street preacher *is* in an *acting* category all its own . : -) without a *believable* performance between them , all we can do is sit back and *watch* the atmosphere , which is *pretty good* in places . The vr sequences *are* way *cool* , but the physical fx -- *such* as miniatures and mattes -- leave a lot to *be desired* . *Watch* out for those *bad* blue-screens

We would n't *mind* a minute of johnny mnemonic if the action *played better* . Too *bad* the debut director *is* n't very *strong* in this department . His *big* finale *is* a sloppy , *silly* mess that runs twenty minutes too long , which *is* way past the *time* that most of our " wet - wired " processors have *already* shut *down* .

Bottom line : yatf (*yet* another *tortured* future) . *Skip* it .

EVALUATION (5)

We 're back in blade runner territory with this one , *conceptual* artist robert longo 's *vision* of a william *gibson-inspired* future where information *is* the commodity to *kill* for . Front and center *is* johnny (keanu reeves) , a " cyber-courier " who *smuggles* data via a " wet-wired " implant . He 's ready to quit the biz and *get* a portion of his long-term *memory* restored , but , first , he has to finish one *last* , *dangerous* job .

The pressing *problem* in johnny mnemonic *is* that keanu reeves seems to have *forgotten* how to play an action hero since his stint on speed . He 's walking wood in a forest of stiffs that includes henry rollins , ice-t , and dina meyer . (dolph lundgren 's street preacher *is* in an *acting* category all its own . : -) without a *believable* performance between them , all we can do is sit back and *watch* the atmosphere , which is *pretty good* in places . The vr sequences *are* way *cool* , but the physical fx -- *such* as miniatures and mattes -- leave a lot to *be desired* . *Watch* out for those *bad* blue-screens

We would n't *mind* a minute of johnny mnemonic if the action *played better* . Too *bad* the debut director *is* n't very *strong* in this department . His *big* finale *is* a sloppy , *silly* mess that runs twenty minutes too long , which *is* way past the *time* that most of our " wet - wired " processors have *already* shut *down* .

Bottom line : yatf (*yet* another *tortured* future) . *Skip* it .

CONCLUSIONS

- Sentiment analysis can benefit from a deep analysis of a text's rhetorical structure, as opposed to a shallow RST analysis or none at all

CONCLUSIONS

- Sentiment analysis can benefit from a deep analysis of a text's rhetorical structure, as opposed to a shallow RST analysis or none at all
- RST-guided polarity classification works best when exploiting RST trees of smaller units of a text

CONCLUSIONS

- Sentiment analysis can benefit from a deep analysis of a text's rhetorical structure, as opposed to a shallow RST analysis or none at all
- RST-guided polarity classification works best when exploiting RST trees of smaller units of a text
- Both nuclei and satellites appear to play an important role in conveying sentiment, whereas satellites have until now been deemed predominantly irrelevant

CONCLUSIONS

- Sentiment analysis can benefit from a deep analysis of a text's rhetorical structure, as opposed to a shallow RST analysis or none at all
- RST-guided polarity classification works best when exploiting RST trees of smaller units of a text
- Both nuclei and satellites appear to play an important role in conveying sentiment, whereas satellites have until now been deemed predominantly irrelevant
- Significantly improved polarity classification performance w.r.t. not accounting for structural aspects of content comes at a cost of increased processing times

FUTURE WORK

- Explore other (faster) methods of identifying discourse structure in natural language text
- Investigate our findings' applicability to vector-based machine learning approaches to sentiment analysis
- Evaluate our findings on different corpora

QUESTIONS?

Alexander Hogenboom
Econometric Institute
Erasmus University Rotterdam
P.O. Box 1738, NL-3000 DR
Rotterdam, the Netherlands
hogenboom@ese.eur.nl