Recognizing Contradictions in Text Mining

Sanda Harabagiu
Human Language Technology Research Institute
University of Texas at Dallas

The problem: Detecting incompatible information in texts

- The premise: Contradictions occur whenever information communicated in two different texts is incompatible.
- Rely on three forms of linguistic information:
  1. Negation
  2. Antonymy
  3. Semantic and pragmatic information associated with CONTRAST discourse relations

Contradictions - examples

NEGATION

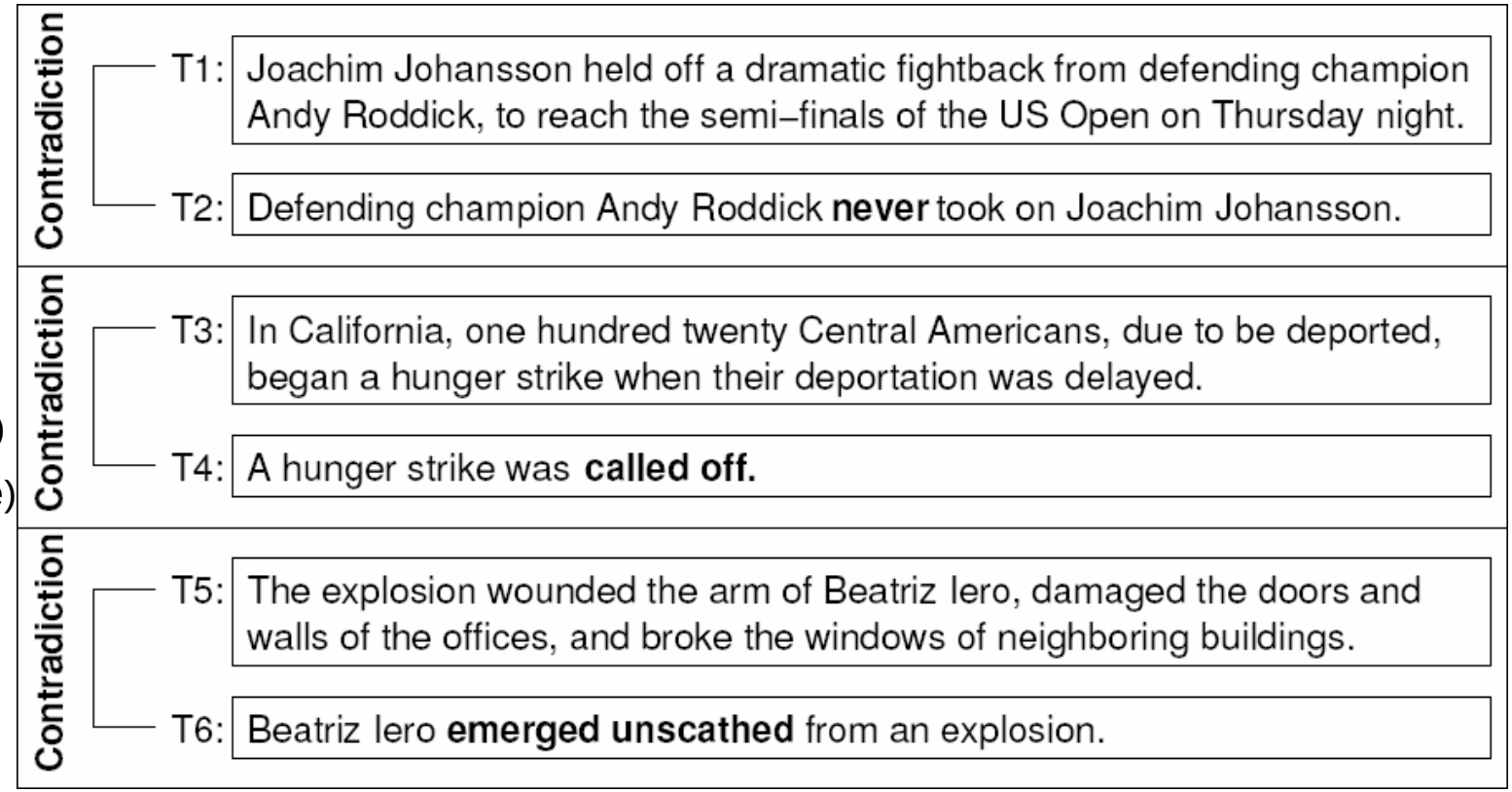
never

ANTONYMS

Begun (hunger strike)

Call off (hunger strike)

CONTRAST



Detecting Incompatible Information in Texts

- Answer Validation (QA) – a special case
  - Filter spurious answers
  - Know when no answer can be returned automatically (without additional processing or knowledge)

Question: When did Pakistan test its Shaheen-2 ballistic missile?
Answer$_1$: The source noted that the Shaheen-2, with a range of 2400 km, has **never** been tested by Pakistan.
Answer$_2$: Pakistan has said that it performed several tests of its 2300 km-range Shaheen-2 missile **in September 2004**.

Detection of Compatible Information

Much work in NLP focused on the detection of *compatible* information:

- Semantic Similarity
  - (Pedersen et al. 2004) WordNet::Similarity
- Redundancy
  - (Carbonell and Goldstein) Maximal Marginal Relevance (MMR)
- Textual Entailment
  - (Glickman and Dagan, 2005) PASCAL Recognizing Textual Entaiment (RTE) Challenge

Textual Entailment

- The RTE task requires systems to recognize when the meaning of a *hypothesis* can be reasonably inferred from the meaning of a *text*.

Text

The Bills now appear ready to hand the reins over to one of their two-top picks from a year ago in quarterback J.P. Losman, who missed most of last season with a broken leg.

Hypothesis

The Bills plan to give the starting job to J.P. Losman.

Textual Entailment

- Systems participating in the 2005 and 2006 PASCAL RTE challenges have made use of a wide range of tools for gauging the compatibility of pairs of texts:
  - Lexico-Semantic resources (NER, semantic parsers)
  - Lexical alignment (must identify portions of text that contain corresponding information) (Raina et al. 2005)
  - Lexical or phrase-level alternations (paraphrases) (Hickl et al. 2006)
  - Syntactic knowledge (need to know which syntactic contexts entailment cannot occur in) (Vanderwende et al. 2005, 2006)
  - Semantic and pragmatic knowledge (factivity, belief, speech act, etc.) (Bobrow et al. 2005)
  - Axioms (Tatu et al. 2005)

Overview

- The problem: Detecting incompatible information in texts
- Framework for recognizing contradictions
  - View 1: remove negations; treat as entailment
  - View 2: treat as a separate phenomenon
- Processing Negation in Text
- Recognizing Contrasts
- Antonymy Detection
- Evaluation
- Conclusions

Recognizing Contradictions

Could an approach similar to TE be used to detect types of incompatible information?

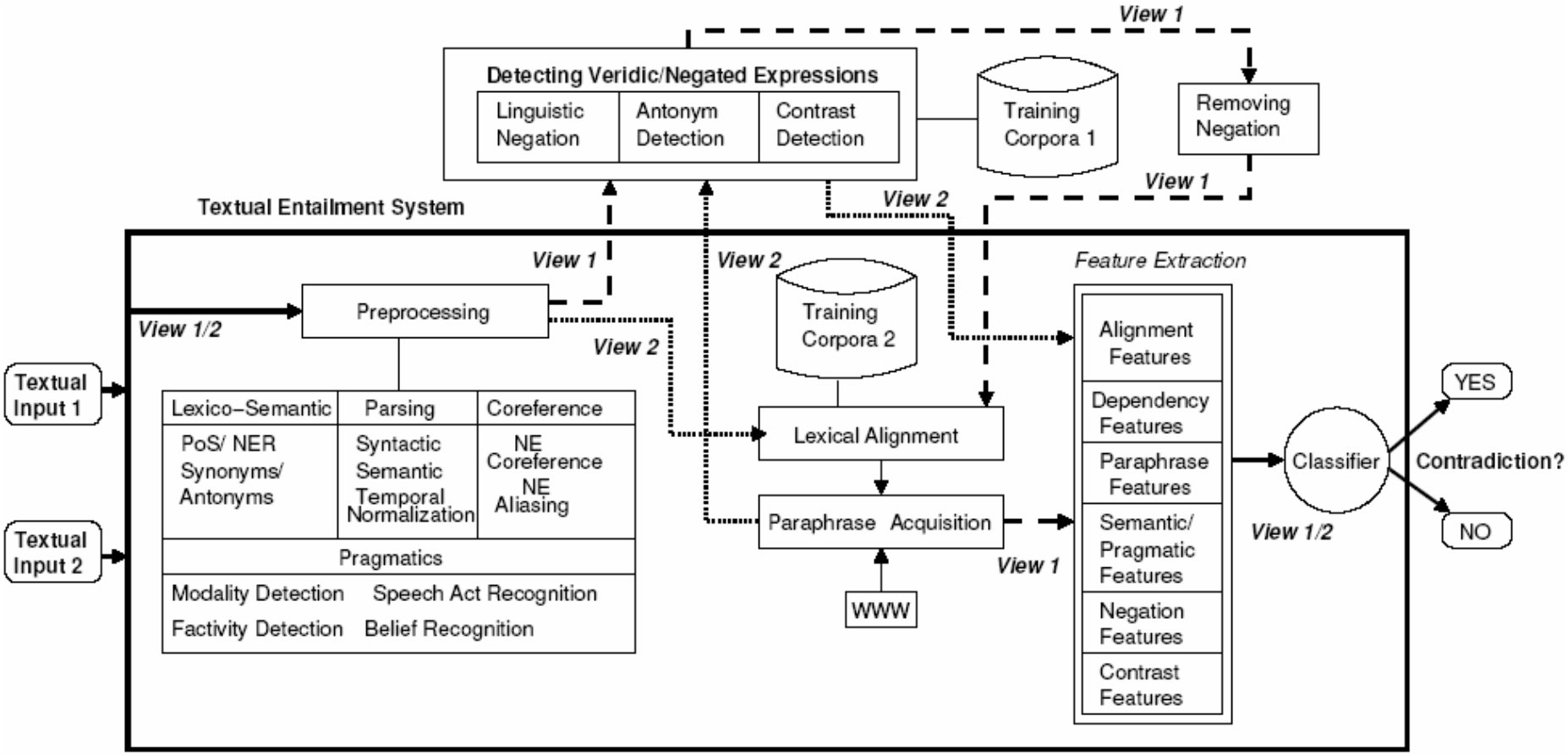
What would be needed?

1. Ability to detect negation
2. Ability to detect instances of antonymy
3. Ability to detect contrasting pairs of sentences
4. Alignment of sentences so we can determine what is incompatible
5. Framework for combining all annotations in order to make a decision.

Two Views on Recognizing Contradictions

- **View 1**: Contradictions are recognized by identifying and removing negations of propositions and then testing for textual entailment
- **View 2**: Contradictions are recognized by
  - *deriving linguistic information* from text inputs, including information that identifies
    - negations
    - contrasts
    - oppositions (antonyms)
  - and by *training a classifier* in a similar way to the classifier for entailment.

Framework for Recognizing Contradictions



Overview

- Detecting incompatible information in texts
- Framework for recognizing contradictions
  - View 1: remove negations; treat as entailment
  - View 2: treat as a separate phenomenon
- Processing Negation in Text
- Recognizing Contrasts
- Antonymy Detection
- Evaluation
- Conclusions

Processing Negation in Text

- Overt Negation
  - Overt negative markers (*n't* and *not*)
  - Negative quantifiers (*no*, *no one*, *nothing*)
  - Strong negative adverbs (*never*)
- Indirectly Licensed Negation
  - Verbs (*deny*, *fail*, *refuse*, *keep from*)
  - Prepositions (*without*, *except*)
  - Weak quantifiers (*few*, *any*, *some*)
  - Traditional negative polarity items (*a red cent*, *any more)*

Detecting Negation in Text

- Step 1: Preprocessing + flag overt and implicit markers of negation
- Step 2: Detect negated events
  - Events are marked as negated if they fall within the syntactic scope of a negative marker
- Step 3: Detect negated entities
  - Any noun phrase that falls within the scope of an overt negative quantifier (e.g. *no*) or a non-veridical quantifier (e.g. *few*, *some*, *any*)
- Step 4: Detect negated states
  - We constructed an ontology of state-denoting terms from WordNet (*symptom*, *condition*, *situation*)
  - A negated state is marked as a state-denoting noun phrase that falls within the scope of a negative marker

Answering Negative Questions

- We evaluated the performance of our negation detection algorithm in the context of a question answering system (Harabagiu et al. 2005).

- Questions that contain negated elements seek sets of entities or events that correspond to the semantic category of an expected answer type and satisfy the negated for of the proposition expressed by the question.

Q_1: What countries have **not** won the World Cup?
Q_2: Which Indonesian islands are **not** in the Southern Hemisphere?
Q_3: What US Army vehicles are **not** personnel transports?
Q_4: What nerve agents does Russia have **other than** VX gas?

Techniques of Answering Negative Questions

1. Detection of negation.
   - Answers are returned as the set of candidate answers that were found in contexts where the proposition found in the question was deemed to be false.
2. Split the question into:
   1) A positive analog of the original negative question
   2) A question that seeks the full extension of the original answer type term.
3. Use textual entailment
   - Retrieve a set of candidate answer passages which feature keywords extracted from the negative question
   - Transform negative questions into declarative statements
   - Answers that are deemed to be entailed by the transformed question are returned as answers to the original question

Evaluation of Answering Negative Questions

Strategy	Accuracy	MRR
Negation Detection	24.7%	0.298
Set Intersection	17.3%	0.213
Entailment	48.0%	0.404
Hybrid	55.3%	0.495

Recognizing Contrasts

- Previous work (Marcu and Echihabi, 2002) relies mainly on cue phrases for the detection of Contrast discourse relations.
- We argue that the presence of opposing information contributes more to the assessment of a CONTRAST discourse relation than the presence of a cue phrase

Such standards would *preclude* arms sales to states like Libya, which is also currently subject to a U.N. embargo.

But states like Rwanda, before its present crisis would still *be able* to legally buy arms.

The mechanisms used for Recognizing CONTRASTs

- Thematic alignment based on PAS trained on Propbank and NomBank
- Five additional forms of alignment between Predicate-Argument Structures:

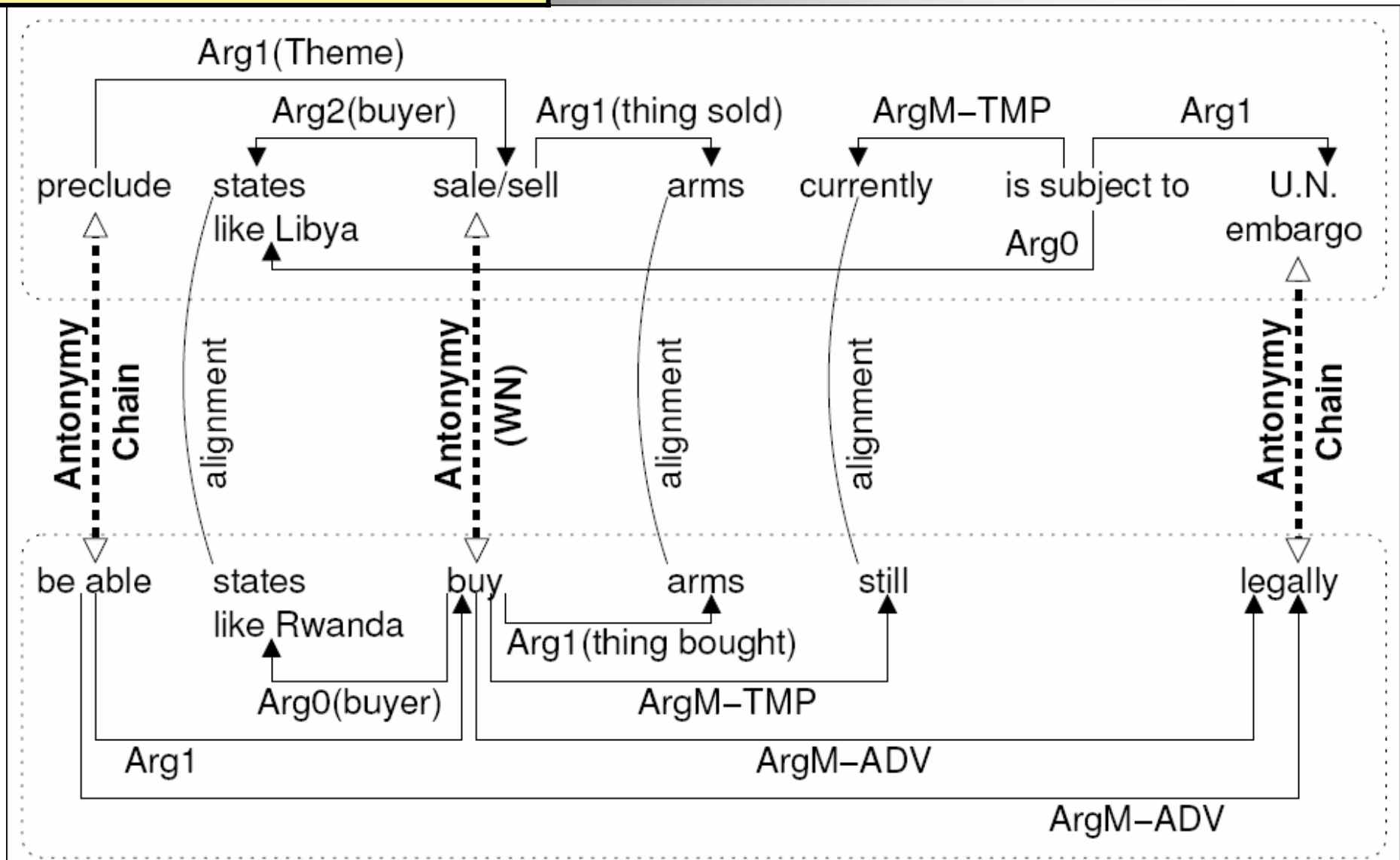
1. Alignments between arguments having identical heads
2. Antonymy relations encoded in WordNet
3. Antonymy chains derived from WordNet for approximating opposing information
4. Alignment based on functional similarity or inclusion
5. Synonym chains

Such standards would *preclude* arms sales to states like Libya, which is also currently subject to a U.N. embargo.

But states like Rwanda, before its present crisis would still *be able* to legally buy arms.

Human Language Technology Research Institute

Example of Alignment



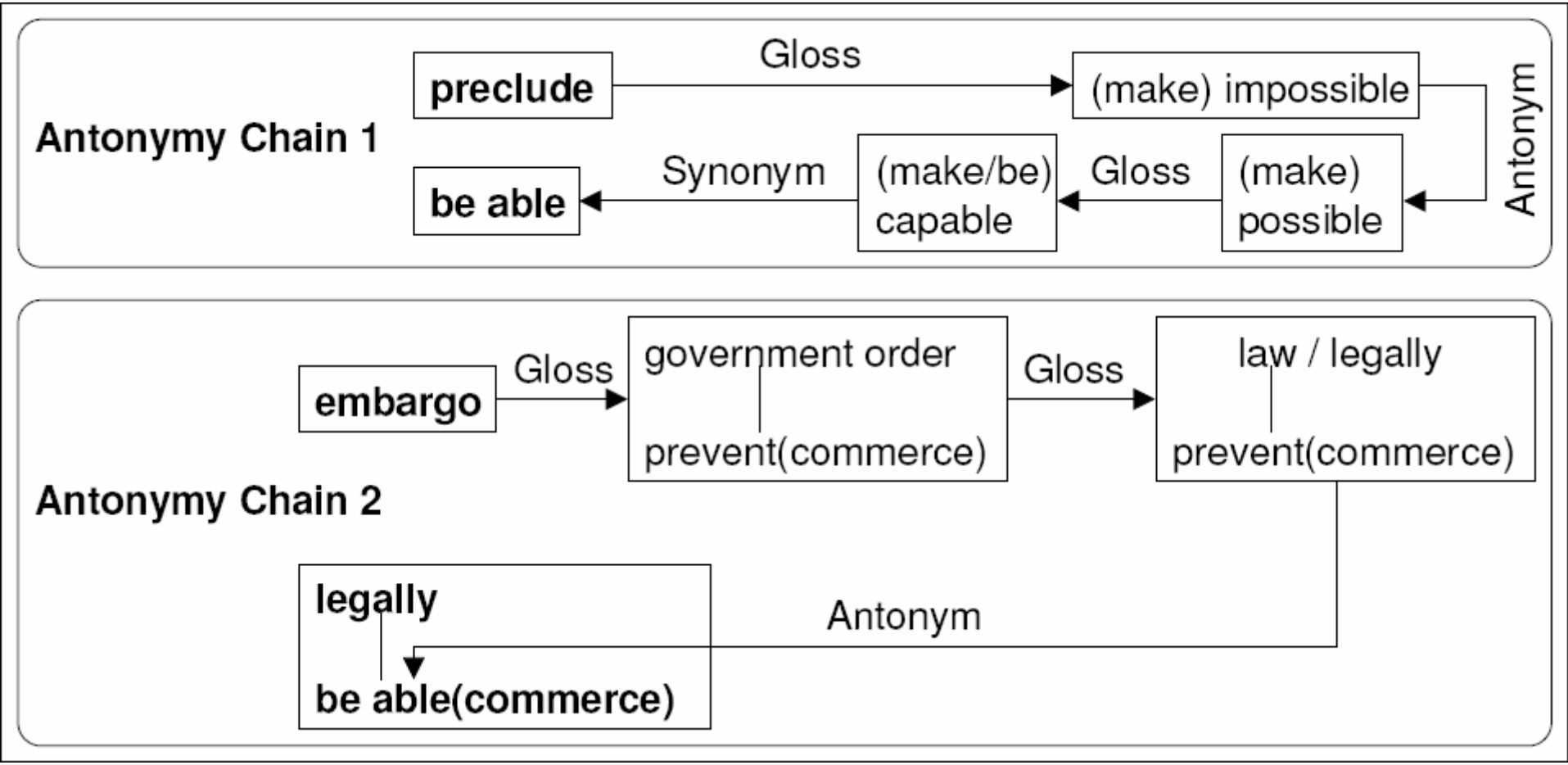
Discovering CONTRAST Discourse Relations

- Step 1: Generate predicate-argument structures
- Step 2: Find possible alignments between predicates and arguments
- Step 3: Acquire paraphrases of aligned information
- Step 4: Cluster paraphrases into sets conveying the same concept
- Step 5: Compute the alignment probability and alignment features
- Step 6: Compute paraphrase features, dependency features, contrast features and semantic/pragmatic features
- Step 7: Use a Maximum Entropy classifier to decide the recognition of a CONTRAST relation

Generating Antonymy Chains

- Step 1: Given a pair of words (w_1, w_2) we retrieve all antonyms encoded in WordNet
- Step 2: By using only Gloss and Is-A relations find lexical chains from each of the words w_i to one of the arguments of the antonymy relation
- Step 3: Limit the length of the chain to the antonymy arguments to 3
- Step 4: Within a gloss, if a dependency relation is considered, a dependency node is created and the dependency is propagated throughout the chain
- Step 5: When both words are connected to the arguments of the antonyms, a chain is established.

Examples of Antonymy Chains



Evaluation

- Recognition of Antonymy
- Identifying CONTRAST relations
- Recognizing Textual Contradiction

Recognition of Antonymy

- 1410 pairs of antonyms extracted from *wikitionary.org*

	Examples	*n*=5	*n*=8	*n*=10
0 antonyms in WN	670	0.1791	0.3134	0.3284
1 antonym in WN	460	0.2174	0.4706	0.5000
2 antonyms in WN	280	0.5185	0.5833	0.6522
TOTAL	1410	0.2589	0.4184	0.4489

Identifying CONTRAST discourse relations

- 10,000 instances of the CONTRAST discourse relation extracted from the web and newswire documents.
- Considered pairs of adjacent sentences featuring the cue phrase “but”.
- 9000 pairs were used o train the system, the remaining 1000 were hand-validated by human annotators and held out as evaluation set.
- Negative examples for CONTRAST were produced by selecting non-adjacent pairs of sentences from each of the documents containiging one positive example.

	Test Size	Accuracy	Avg. Precision
Contrast	1200	74.19%	74.00%

Textual Contradiction

- Used the data prepared for the 2006 PASCAL Recognizing Textual Entailment (RTE) Challenge:
  - Negate one of the sentences in each of the example pairs
  - Create paraphrases of each of the negated sentences

	Text Size	Accuracy	Avg. Precision
Textual Entailment			
PASCAL RTE 2006	800	75.38%	80.16%
Contradiction			
Negation Only (View 1)	800	75.63%	68.07%
Paraphrase Only (View 2)	800	62.55%	67.35%
Paraphrase+Negation (V2)	800	64.00%	75.74%

Conclusions

- New framework for recognizing contradictions in texts.
- The approach that treats contradictions as a separate phenomenon (View 2) proves to be better; it depends on significant enhancements to the state-of-the-art in:
  - Negation Processing
  - Contrast Detection
  - Antonymy Detection
- This combination of techniques leads to identifying contradictions with over 62% accuracy.