Text to Text Generation

A Candidacy Exam

Kapil Thadani

Columbia University
New York, NY

DECEMBER 2009

Text to Text Generation

- Rewriting of text according to requirements
- Potential operations:
  - Rewording and rearranging phrases
  - Combining or splitting up sentences
  - Deleting content
- Applications: summarization/redundancy, question-answering

Text to Text Generation

1. **Paraphrase induction**
   - Rewording and rearranging phrases
   - Preserve original meaning

Ms. Palin supported the bridge project while running for governor, and abandoned it after it became a national scandal.

↓

After it became a national scandal, Ms. Palin abandoned the bridge project that she had supported during her gubernatorial campaign.

Text to Text Generation

1. **Paraphrase induction**
    - Rewording and rearranging phrases
    - Preserve original meaning

Ms. Palin supported the bridge project while running for governor, and abandoned it after it became a national scandal.

↓

After it became a national scandal, Ms. Palin abandoned the bridge project that she had supported during her gubernatorial campaign.

Text to Text Generation

1 Paraphrase induction

- Rewording and rearranging phrases
- Preserve original meaning

Ms. Palin supported the bridge project while running for governor, and abandoned it <u>after it became a national scandal</u>.

↓

<u>After it became a national scandal</u>, Ms. Palin abandoned the bridge project that she had supported during her gubernatorial campaign.

Text to Text Generation

2 Sentence simplification

- Splitting up sentences for readability
- Preserve original meaning

Ms. Palin supported the bridge project while running for governor, and abandoned it after it became a national scandal.

↓

Ms. Palin supported the bridge project while running for governor.

She abandoned it after it became a national scandal.

Text to Text Generation

3 Sentence compression

- Delete content for summarization
- Preserve *important* aspects

Ms. Palin supported the bridge project while running for governor, and abandoned it after it became a national scandal.

$\downarrow$

Ms. Palin abandoned the bridge project after it became a national scandal.

Text to Text Generation

4 Sentence fusion

- Combine sentences for summarization
- Preserve *important* aspects

Ms. Palin supported the bridge project while running for governor, and abandoned it after it became a national scandal.

The media keeps repeating that Palin actually turned against the bridge project only after it became a national symbol of wasteful spending.

↓

Ms. Palin supported the bridge project while running for governor, and abandoned it only after it became a became a national symbol of wasteful spending.

Text to Text Generation

4 Sentence fusion

- Combine sentences for summarization
- Preserve *important* aspects

Ms. Palin supported the bridge project while running for governor, and abandoned it after it became a national scandal.

The media keeps repeating that Palin actually turned against the bridge project <u>only after it became a national symbol of wasteful spending</u>.

↓

Ms. Palin supported the bridge project while running for governor, and abandoned it <u>only after it became a became a national symbol of wasteful spending</u>.

Text to Text Generation

Dimensions

1. Preservation of sentence semantics:
   - Lossless (paraphrasing, simplification)
   - Lossy (compression, fusion)

2. Transformation of sentences:
   - One to one (paraphrasing, compression)
   - One to many (simplification)
   - Many to one (fusion)

Text to Text Generation

Reluctant paraphrase

Dras (1997)

- Links paraphrasing, simplification and compression
- Mathematical optimization with constraints
  1. Sentence length
  2. Readability
  3. Lexical density
- Paraphrasing only carried out if a constraint is violated

M. Dras (1997) Reluctant Paraphrase: Textual Restructuring under an Optimisation Model

Text to Text Generation

Reluctant paraphrase

Dras (1997)

- Links paraphrasing, simplification and compression
- Mathematical optimization with constraints
  1. Sentence length
     - Number of words ↓
  2. Readability
  3. Lexical density
- Paraphrasing only carried out if a constraint is violated

M. Dras (1997) Reluctant Paraphrase: Textual Restructuring under an Optimisation Model

Text to Text Generation

Reluctant paraphrase

Dras (1997)

- Links paraphrasing, simplification and compression
- Mathematical optimization with constraints
  1. Sentence length
     - Number of words $\downarrow$
  2. Readability
     - Sentence complexity (average sentence length) $\downarrow$
     - Word complexity (average number of syllables OR proportion of infrequent words) $\downarrow$
  3. Lexical density
- Paraphrasing only carried out if a constraint is violated

M. Dras (1997) Reluctant Paraphrase: Textual Restructuring under an Optimisation Model

Text to Text Generation

Reluctant paraphrase

Dras (1997)

- Links paraphrasing, simplification and compression
- Mathematical optimization with constraints
  1. Sentence length
     - Number of words $\downarrow$
  2. Readability
     - Sentence complexity (average sentence length) $\downarrow$
     - Word complexity (average number of syllables OR proportion of infrequent words) $\downarrow$
  3. Lexical density
     - Proportion of non-content words $\leftrightarrow$
- Paraphrasing only carried out if a constraint is violated

M. Dras (1997) Reluctant Paraphrase: Textual Restructuring under an Optimisation Model

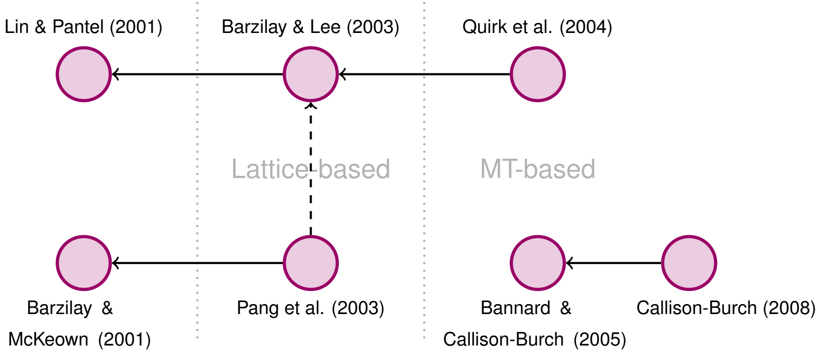
Paraphrase Induction

Two strategies

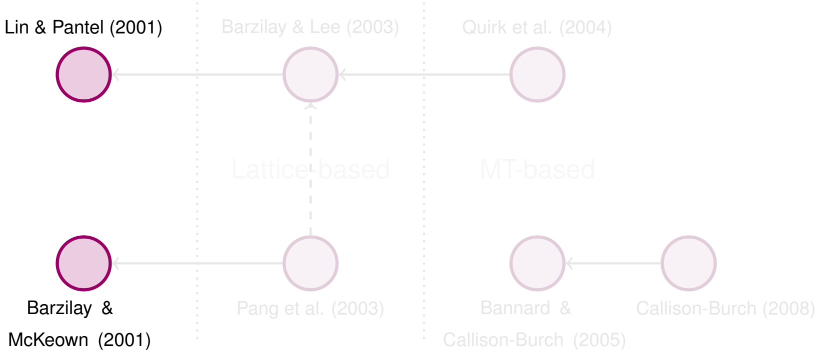
1. **Non-parallel corpora**
   - Easier to obtain data
   - Harder to detect accurate paraphrases

2. **Parallel corpora**
   - Harder to obtain data
   - Easier to detect accurate paraphrases

Paraphrase Induction



Paraphrase Induction



Learning paraphrase rules

Paraphrase Induction

Learning paraphrase rules

Lin & Pantel (2001)

1. Standard (non-parallel) corpus: newswire
2. Extended distributional hypothesis:
   - Extract slotted paths from dependency trees
   - Paths which tend to have similar slots are similar

Barzilay & McKeown (2001)

1. Parallel corpora: multiple translations of French novels
2. Bootstrap through co-training:
   - Identify predictive context patterns around paraphrases
   - Identify paraphrases within predictive context patterns

D. Lin & P. Pantel (2001) DIRT - Discovery of Inference Rules in Text

R. Barzilay & K. McKeown (2001) Extracting Paraphrases from Parallel Corpora

Paraphrase Induction

Learning paraphrase rules

Lin & Pantel (2001)

Inference rules:

- NN_0 is the author of $NN_1 \equiv NN_0$ wrote NN_1
- NN_0 solved $NN_1 \equiv NN_0$ found a solution to NN_1
- NN_0 caused $NN_1 \equiv NN_1$ is triggered by NN_0

Barzilay & McKeown (2001)

Lexical paraphrases:

- burst into tears $\equiv$ cried
- comfort $\equiv$ console
- countless $\equiv$ lots of

Morpho-syntactic patterns:

- NN_0 POS $NN_1 \equiv NN_1$ IN DET NN_0
  King's son son of the King
- VB_0 to $VB^1 \equiv VB_0\ VB_1$
  start to talk start talking

D. Lin & P. Pantel (2001) DIRT - Discovery of Inference Rules in Text

R. Barzilay & K. McKeown (2001) Extracting Paraphrases from Parallel Corpora

Paraphrase Induction

Learning paraphrase rules

Lin & Pantel (2001)

3 Human evaluation on paths from 6 TREC-8 questions:
- Accuracy varies from 35% - 92.5%
- Very low overlap with manually generated paraphrases

4 Can't distinguish between synonymy and antonymy!

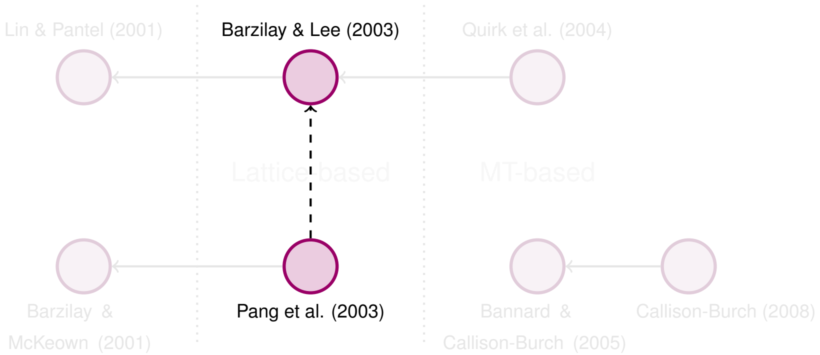
Barzilay & McKeown (2001)

3 Human evaluation of generated paraphrases:
- 86.5% out of context; 91.6% in context
- 69% overlap with human paraphrases in small recall study

4 65% of paraphrases extend beyond synonymy

D. Lin & P. Pantel (2001) DIRT - Discovery of Inference Rules in Text
R. Barzilay & K. McKeown (2001) Extracting Paraphrases from Parallel Corpora

Paraphrase Induction



Lattice-based methods

Paraphrase Induction

Lattice-based methods

Barzilay & Lee (2003)

1. Comparable corpora: clustered newswire articles
2. Multiple sequence alignment (MSA) over clustered sentences
    - Regions of high variability are *slots*
    - Lattices from sentences with similar slot values are similar

Pang et al. (2003)

1. Parallel corpus: Multiple Translation Chinese
2. FSAs from parse forests created by merging syntactic trees
    - *Squeezing* to remove redundancy
    - Different FSA paths are paraphrases

R. Barzilay & L. Lee (2003) Learning to Paraphrase: An Unsupervised Approach using Multiple-Sequence Alignment

B. Pang, K. Knight & D. Marcu (2003) Syntax-based Alignment of Multiple Translations: Extracting Paraphrases and Generating New Sentences

Paraphrase Induction

Lattice-based methods

Barzilay & Lee (2003)

- Human evaluation on AFP/Reuters:
  - 38% improvement over templates from Lin & Pantel (2001)
  - Generated paraphrases from 59 sentences judged better than Wordnet substitution baseline

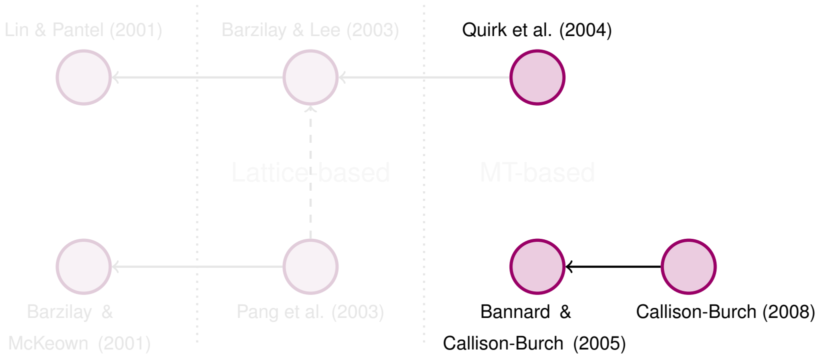
Pang et al. (2003)

- Human evaluation on parallel sentence groups:
  - 15% improvement over Barzilay & McKeown (2001), but half the number of paraphrases generated
  - Word repetition 10x more more likely for MSA algorithm

R. Barzilay & L. Lee (2003) Learning to Paraphrase: An Unsupervised Approach using Multiple-Sequence Alignment

B. Pang, K. Knight & D. Marcu (2003) Syntax-based Alignment of Multiple Translations: Extracting Paraphrases and Generating New Sentences

Paraphrase Induction



MT-based methods

Paraphrase Induction

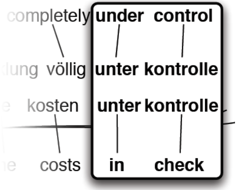
MT-based methods

Quirk et al. (2004)

1. Comparable corpora: online news clusters
2. Phrase-based MT approach; no inter-phrase reordering
3. Human evaluation on 59 sentence corpus
   - Paraphrases better than Barzilay & Lee (2003) and far less information added and lost
   - More general than MSA (15/59 paraphrases from one template)

Bannard & Callison-Burch (2005), Callison-Burch (2008)

1. Bilingual parallel corpora from Europarl
2. Pivot through foreign-language corpus
   - No syntax: equal, created equal, to create equal
3. Manual evaluation



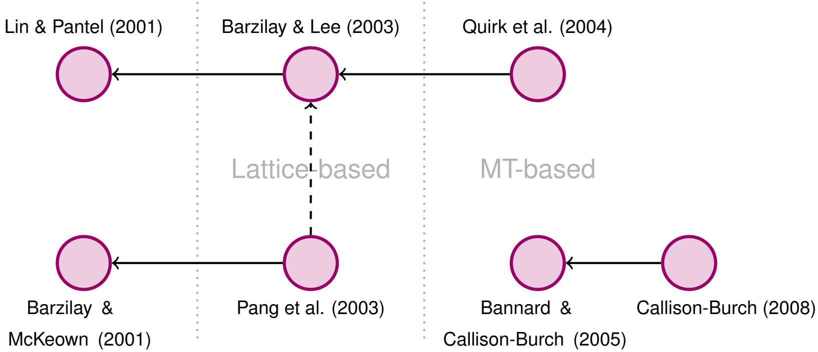
	Meaning	Grammar	Both
No syntax (B&CB '05)	46%	44%	36%
With syntax (CB '08)	61%	68%	55%

C. Quirk, C. Brockett & W. Dolan (2004) Monolingual Machine Translation for Paraphrase Generation

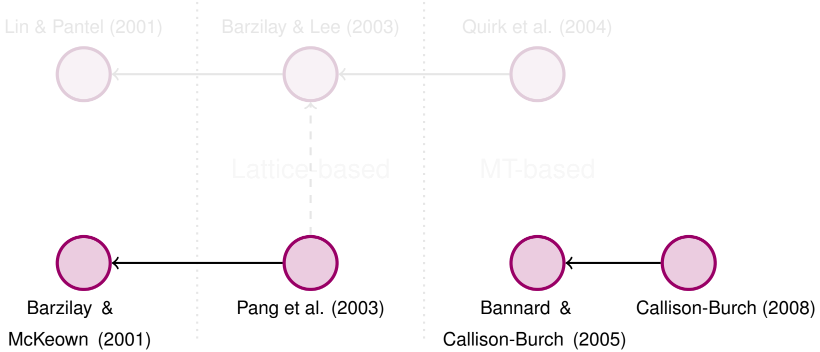
C. Bannard & C. Callison-Burch (2005) Paraphrasing with Bilingual Parallel Corpora

C. Callison-Burch (2008) Syntactic Constraints on Paraphrases Extracted from Parallel Corpora

Paraphrase Induction



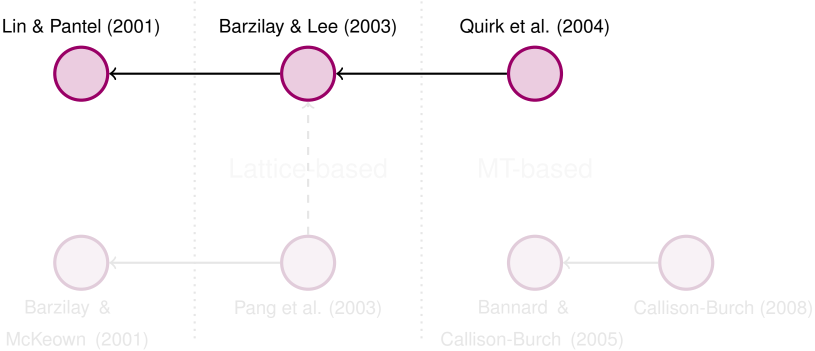
Paraphrase Induction



Corpus strategy

- Parallel

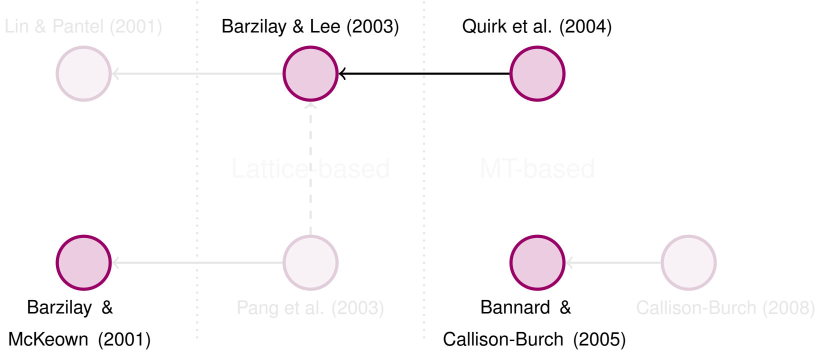
Paraphrase Induction



Corpus strategy

- Parallel
- Non-parallel

Paraphrase Induction



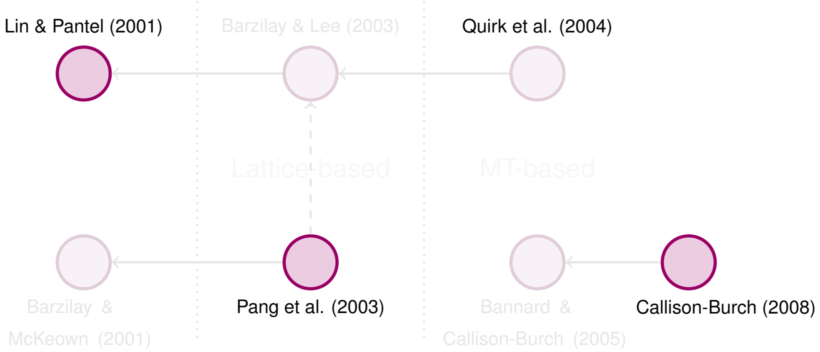
Corpus strategy

- Parallel
- Non-parallel

Features used

- Word/POS

Paraphrase Induction



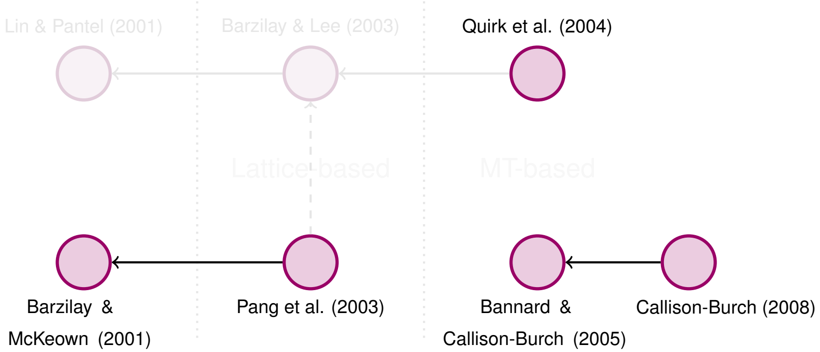
Corpus strategy

- Parallel
- Non-parallel

Features used

- Word/POS
- Syntax

Paraphrase Induction



Corpus strategy

- Parallel
- Non-parallel

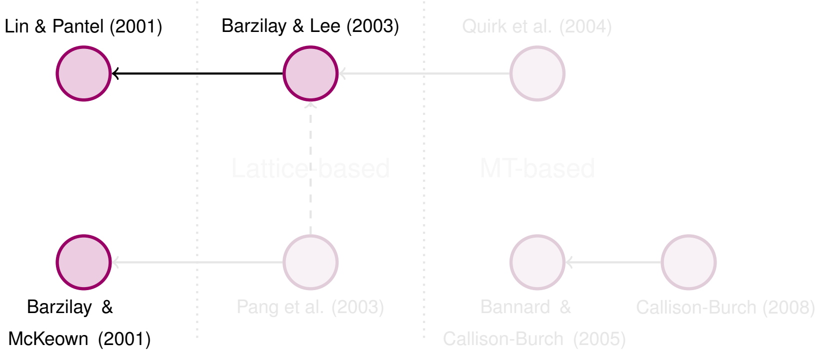
Features used

- Word/POS
- Syntax

Paraphrase type

- Lexical

Paraphrase Induction



Corpus strategy
- Parallel
- Non-parallel

Features used
- Word/POS
- Syntax

Paraphrase type
- Lexical
- Slotted

Sentence Simplification

Klebanov et al. (2004)

Chandrasekar & Bangalore (1997)

Siddharthan (2006)

Enumeration

Hickl (2008)

Sentence Simplification

Chandrasekar & Bangalore (1997)

Siddharthan (2006)

Klebanov et al. (2004)

Enumeration

Hickl (2008)

Simplification rules

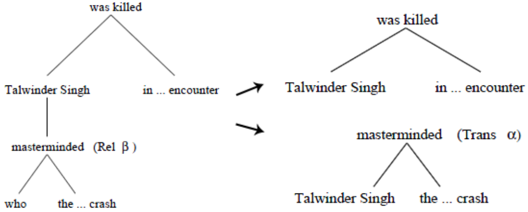
Sentence Simplification

Simplification rules

Chandrasekar & Bangalore (1997)

- Supervised approach based on lexicalized TAG
  - Training parsed using a *lightweight dependency analyzer*
  - Transformation rules generalized from tree pairs

Talwinder Singh, who masterminded the 1984 Kanishka crash, was killed in a fierce two-hour encounter.



Talwinder Singh masterminded the 1984 Kanishka crash.

Talwinder Singh was killed in a fierce two-hour encounter.

- No evaluation!

R. Chandrasekar & S. Bangalore (1997) Automatic Induction of Rules for Text Simplification

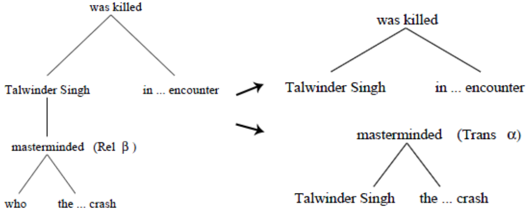
Sentence Simplification

Simplification rules

Chandrasekar & Bangalore (1997)

- Supervised approach based on lexicalized TAG
    - Training parsed using a *lightweight dependency analyzer*
    - Transformation rules generalized from tree pairs

Talwinder Singh, who masterminded the 1984 Kanishka crash, was killed in a fierce two-hour encounter.



Talwinder Singh masterminded the 1984 Kanishka crash.

Talwinder Singh was killed in a fierce two-hour encounter.

- No evaluation!

R. Chandrasekar & S. Bangalore (1997) Automatic Induction of Rules for Text Simplification

Sentence Simplification

Simplification rules

Siddharthan (2006)

- Local ordering through recursive transformation:

  1. Mr. Anthony, who runs an employment agency, decries program trading, but **he** isn't sure **it** should be strictly regulated.

  $\downarrow$

  2. Mr. Anthony, who runs an employment agency, decries program trading.
  But **he** isn't sure **it** should be strictly regulated.

  $\downarrow$

  3. Mr. Anthony runs an employment agency.
  He decries program trading.
  But **he** isn't sure **it** should be strictly regulated.

- Focus on conjunctive cohesion and anaphoric cohesion
- Knowledge-heavy: anaphora resolution, clause/appositive identification and attachment, RST-based sentence analysis
- Human evaluation on 95 news sentences:
  - High (95%) grammaticality and meaning preservation

A. Siddharthan (2006) Syntactic Simplification and Text Cohesion

Sentence Simplification

Klebanov et al. (2004)

Chandrasekar & Bangalore (1997)

Siddharthan (2006)

Enumeration

Hickl (2008)

Enumerating propositions

Sentence Simplification

Enumerating propositions

Harriet Beecher Stowe is a writer. She was born in Litchfield Connecticut, USA, the daughter of Lyman Beecher. Raised by her severe Calvinist father, she was educated and then taught at the Hartford Female Seminary.

↓

Harriet Beecher Stowe is a writer.
Harriet Beecher Stowe was born in Litchfield, Connecticut.
Harriet Beecher Stowe was the daughter of Lyman Beecher.
Harriet Beecher Stowe was raised by her severe Calvinist father.
Harriet Beecher Stowe was raised by Lyman Beecher.
Lyman Beecher is Harriet Beecher Stowe's father.
Harriet Beecher Stowe was educated at the Hartford Female Seminary.
Harriet Beecher Stowe taught at the Hartford Female Seminary.

Sentence Simplification

Enumerating propositions

Klebanov et al. (2004)

1. Easy-access sentences: grammatical, single verb, NEs
2. Useful for *information-seeking* applications like QA
3. Relies on rules over dependency parses from MINIPAR
4. Human evaluation on 123 sentences:
   - 55% accuracy; nearly all errors in information maintenance

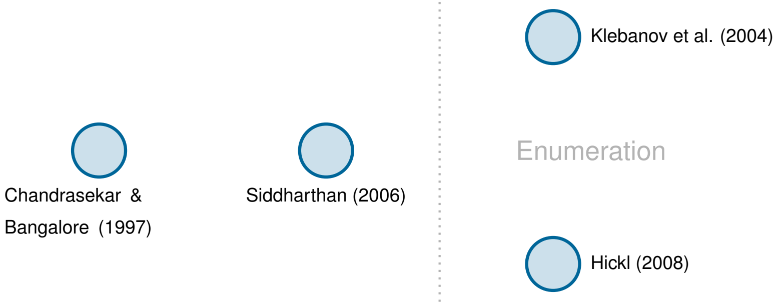
Hickl (2008)

1. Discourse commitments: lightweight propositions that can be inferred as true
2. Applied to textual entailment
3. Uses syntactic and semantic parsing, relation extraction, coreference resolution
4. Entailment evaluation over RTE-2 and RTE-3:
   - 83% correct; state of the art performance by far

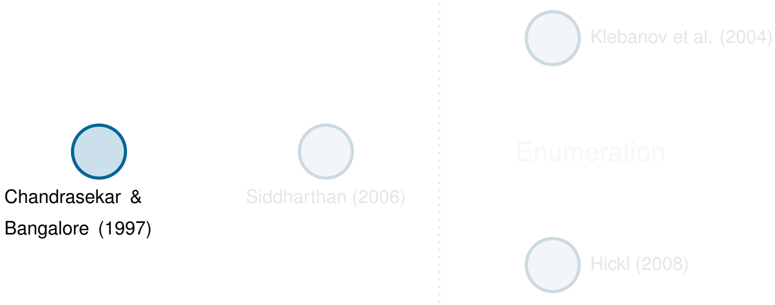
B. Klebanov, K. Knight & D. Marcu (2004) Text Simplification for Information-Seeking Applications

A. Hickl (2008) Using Discourse Commitments to Recognize Textual Entailment

Sentence Simplification

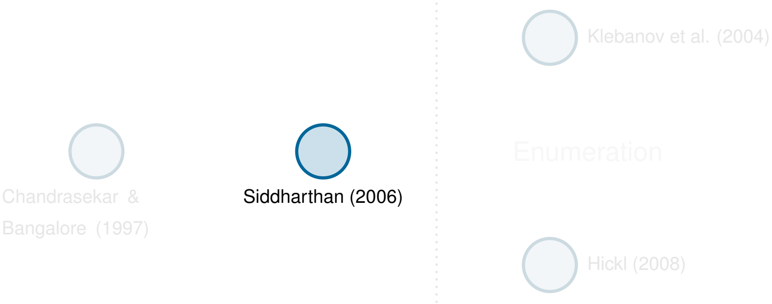


Sentence Simplification



- Just simplification

Sentence Simplification



- Just simplification
- Readability

Sentence Simplification

Klebanov et al. (2004)

Enumeration

Chandrasekar & Bangalore (1997)

Siddharthan (2006)

Hickl (2008)

- Just simplification
- Readability

- Information extraction

Sentence Simplification

Klebanov et al. (2004)

Enumeration

Chandrasekar & Bangalore (1997)

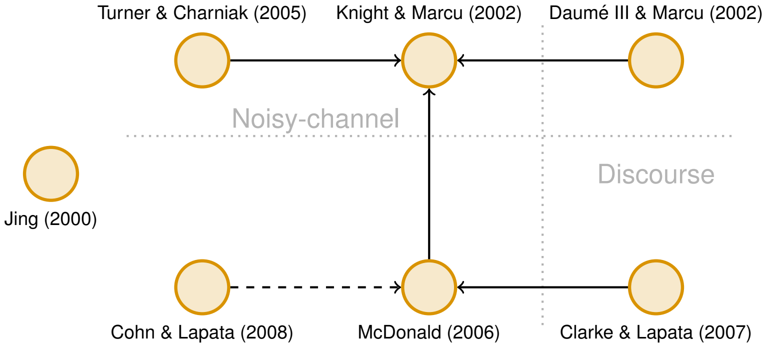
Siddharthan (2006)

Hickl (2008)

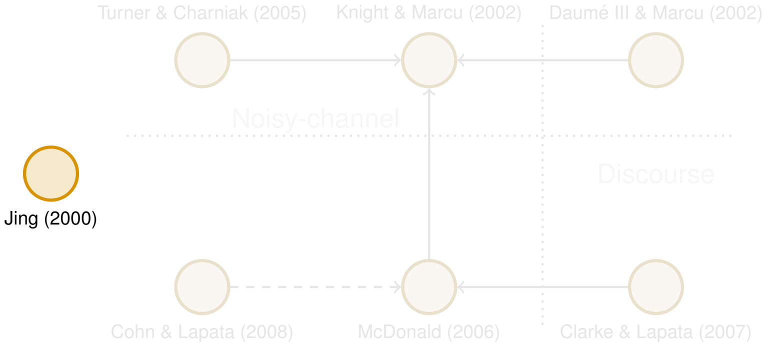
- Just simplification
- Readability

- Information extraction
- Meaning representation

Sentence Compression



Sentence Compression



Sentence reduction

Sentence Compression

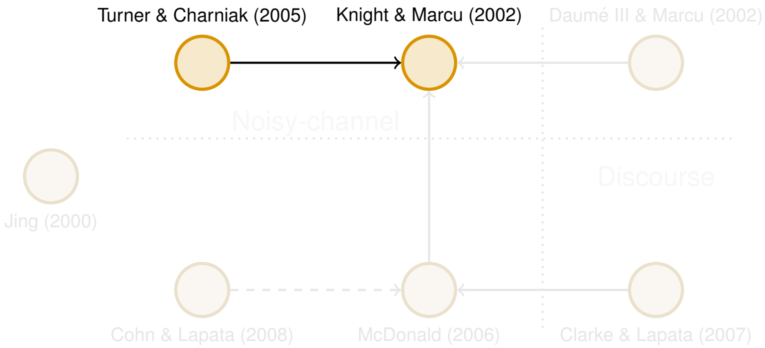
Sentence reduction

Jing (2000)

- Introduced sentence compression for summarization
- Probabilities of dropping particular subtrees estimated using
  - Corpus of human abstracts; sentences aligned to reduced forms
  - Large-scale syntactic lexicon to find obligatory verb arguments
  - Topicality from local context through Wordnet, heuristics
- Evaluation over 100 sentences:
  - Algorithm made same choice as humans 81.3% of the time
  - Compression rate 67% versus 58% for humans

H. Jing (2000) Sentence Reduction for Automatic Text Summarization

Sentence Compression



Noisy-channel models

Sentence Compression

Noisy-channel models

Goal: retrieve compressed **source** string from "noisy" **target** string

	Knight & Marcu (2002)	Turner & Charniak (2005)
Source model	PCFG expansions + bigrams	Syntactic LM
Channel model	Stochastic parse tree rules from aligned ZD corpus	+ unsupervised PCFG expansions from WSJ
Generation	NLG system	Direct
Evaluation:		
- Grammar	4.57	4.79
- Importance	3.85	4.18
- Comp. rate	70.4%	81.2%

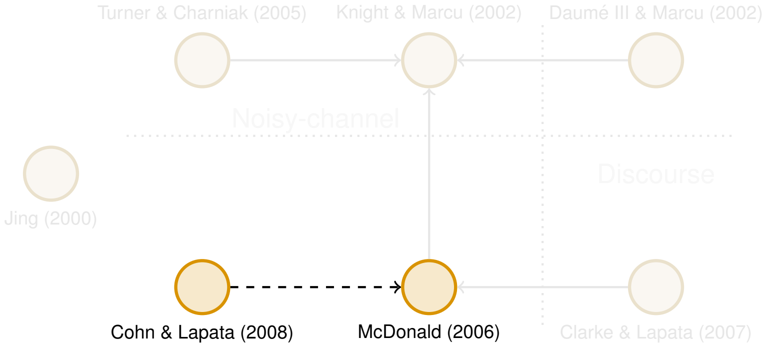
Noisy-channel model not well suited to compression:

- Expansions required to be more likely in the channel model
- Most probable compressed sentence will be almost like the original

K. Knight & D. Marcu (2002) Summarization beyond Sentence Extraction: A Probabilistic Approach to Sentence Compression

J. Turner & E. Charniak (2005) Supervised and Unsupervised Learning for Sentence Compression

Sentence Compression



Discriminative approaches

Sentence Compression

Discriminative approaches

McDonald (2006)

- Uses potentially-noisy parse features as “soft syntactic evidence”
- 78923 features including:
  - POS features for retained & dropped words
  - Dependency features
  - Phrase-structure features for dropped productions
  - ... but no lexical features except dropped verbs/negations
- Human evaluation on Ziff-Davis corpus:
  - Improvements in grammaticality and importance against the decision-tree model in Knight & Marcu (2002)
  - Better importance score than humans, perhaps because of compression rate

R. McDonald (2006) Discriminative Sentence Compression with Soft Syntactic Evidence

K. Knight & D. Marcu (2002) Summarization beyond Sentence Extraction: A Probabilistic Approach to Sentence Compression

Sentence Compression

Discriminative approaches

Cohn & Lapata (2008)

- Allows rewriting, substitutions and insertions
- Synchronous tree substitution grammar (STSG) learns rules for rewriting tree fragments, like Chandrasekar & Bangalore (1997)
  - Compression rules from abstractive corpus of 575 sentences
  - Paraphrasing rules using Bannard & Callison-Burch (2005)
- Grammaticality maintained with language model (LM)
- Human evaluation on abstractive corpus:
  - Gain in importance score over extractive version; implicitly beats McDonald (2006)
  - Proportion of deletions, substitutions and insertions closely mirrors human summarizers

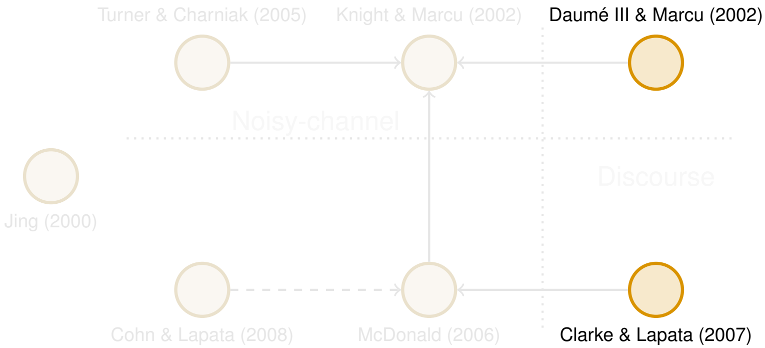
T. Cohn & M. Lapata (2008) Sentence Compression beyond Word Deletion

R. Chandrasekar & S. Bangalore (1997) Automatic Induction of Rules for Text Simplification

C. Bannard & C. Callison-Burch (2005) Paraphrasing with Bilingual Parallel Corpora

R. McDonald (2006) Discriminative Sentence Compression with Soft Syntactic Evidence

Sentence Compression



Considering discourse

Sentence Compression

Considering discourse

Clarke & Lapata (2007)

- Doesn't require aligned corpus of compressed sentences
- Integer linear programming (ILP) with *discourse constraints*
  1. **Centering theory:** single entity salient in an utterance
  2. **Lexical chains:** sequences of related words indicate cohesion across utterances
- Automatic evaluation over manually compressed corpus using grammatical relations:
  - Significantly outperforms McDonald (2006)
- Novel QA evaluation over 6 documents:
  - Significant advantage over McDonald (2006); non-significant difference with gold standard

---

J. Clarke & M. Lapata (2007) Modelling Compression with Discourse Constraints

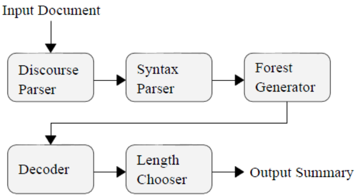
R. McDonald (2006) Discriminative Sentence Compression with Soft Syntactic Evidence

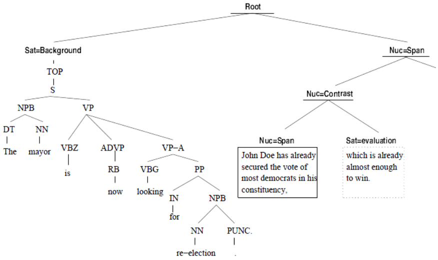
Sentence Compression

Considering discourse

Daumé III & Marcu (2002)

Application of Knight & Marcu (2002) to discourse trees





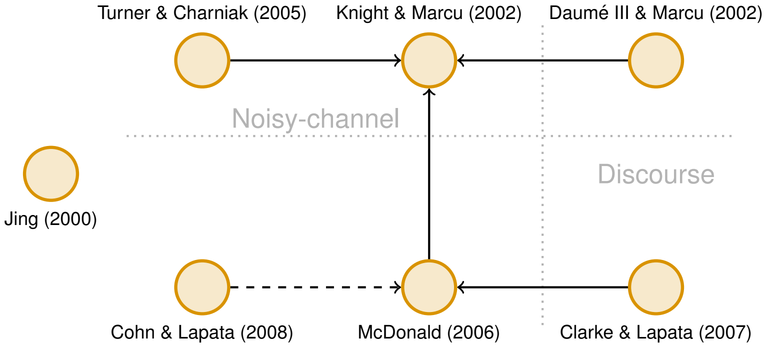
- Compresses *across sentence boundaries*
- Human evaluation over WSJ and small Mitre corpus:
  - Outperforms Knight & Marcu (2002) baseline
  - Better compression rate than Clarke & Lapata (2007)

H. Daumé III & D. Marcu (2002) A Noisy-Channel Model for Document Compression

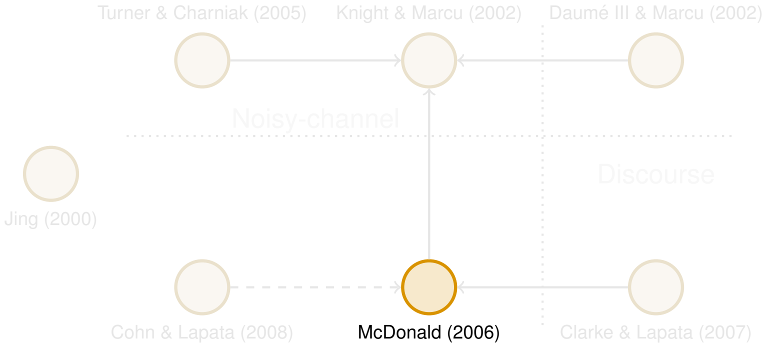
K. Knight & D. Marcu (2002) Summarization beyond Sentence Extraction: A Probabilistic Approach to Sentence Compression

J. Clarke & M. Lapata (2007) Modelling Compression with Discourse Constraints

Sentence Compression

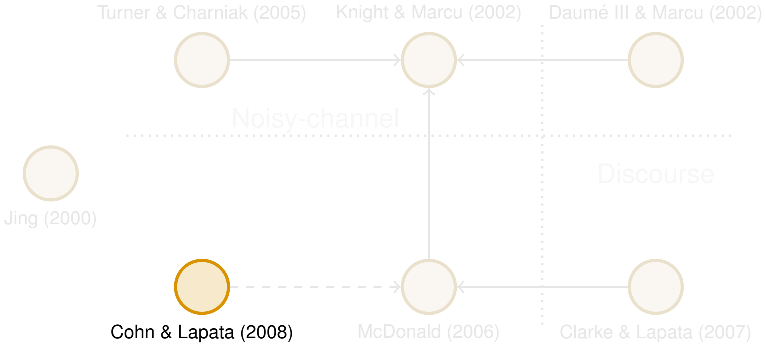


Sentence Compression



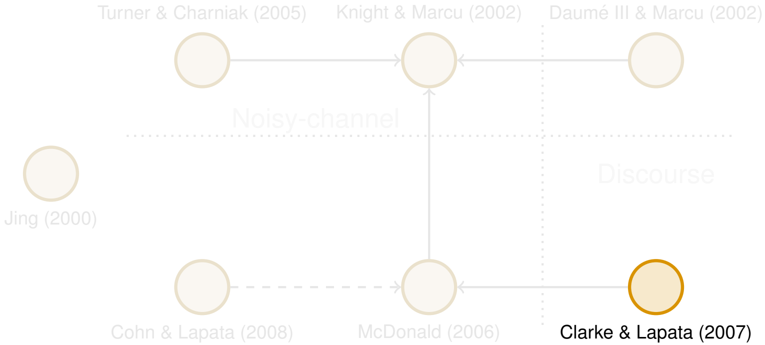
- Handle parse errors

Sentence Compression



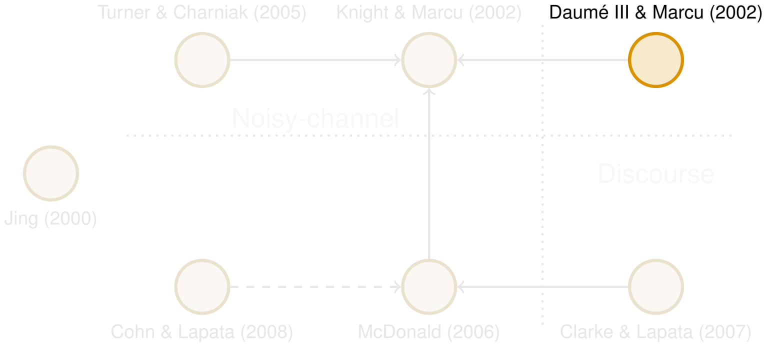
- Handle parse errors
- Rewrite sentences

Sentence Compression



- Handle parse errors
- Rewrite sentences
- No aligned corpus

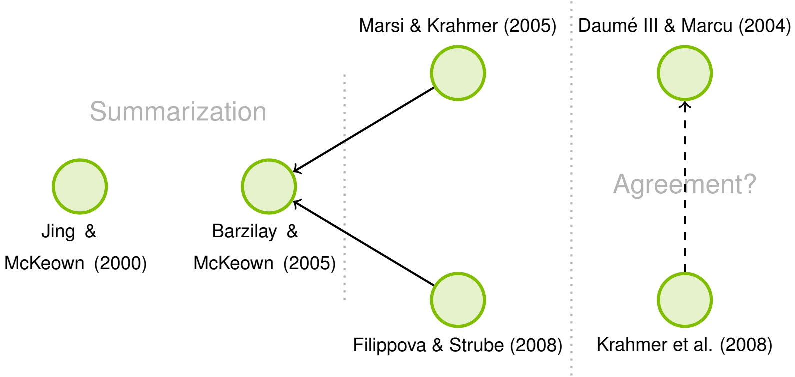
Sentence Compression



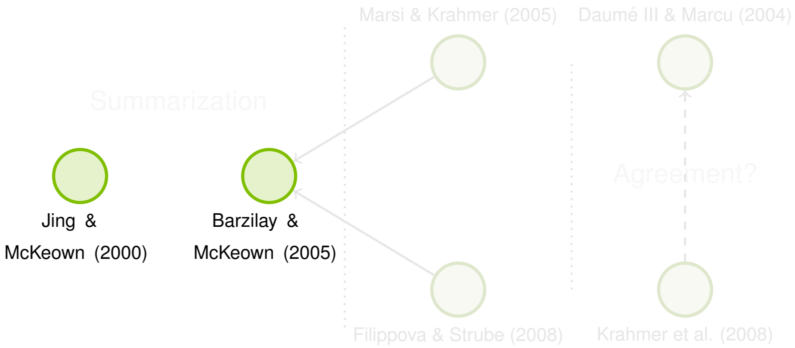
- Handle parse errors
- Rewrite sentences

- No aligned corpus
- Collapse sentences

Sentence Fusion



Sentence Fusion



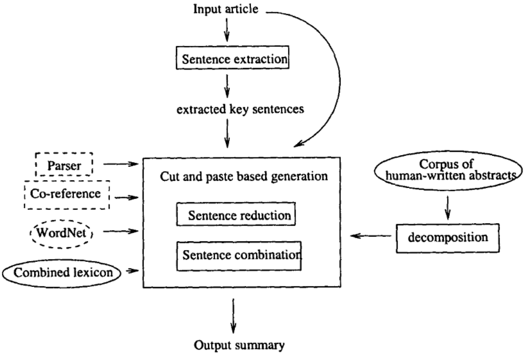
Sentence combination for summarization

Sentence Fusion

Sentence combination for summarization

Jing & McKeown (2000)

1. Summary *revision* using Jing (2000) for reduction



2. Human evaluation on 20 documents:
   - Better conciseness and cohesion than extracted summaries
   - Revised summaries 41% shorter

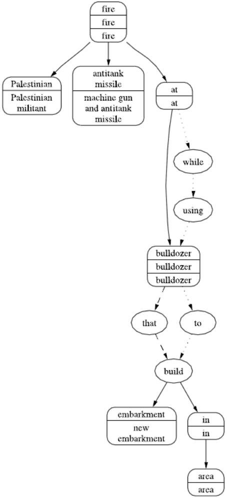
H. Jing & K. McKeown (2000) Cut and Paste Based Text Summarization
H. Jing (2000) Sentence Reduction for Automatic Text Summarization

Sentence Fusion

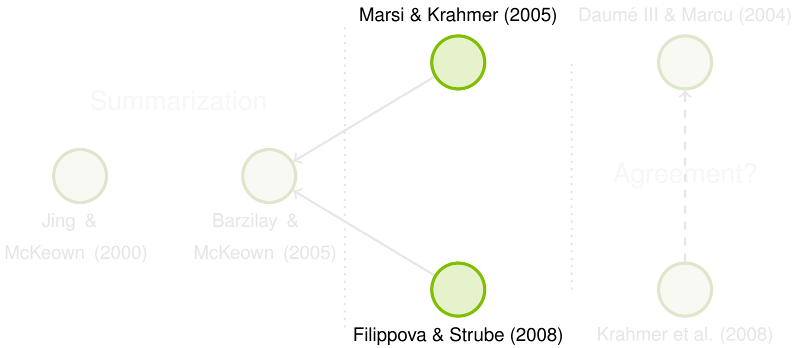
Sentence combination for summarization

Barzilay & McKeown (2005)

- Combination of sentences in a *theme*
- Three components:
  1. Dependency tree alignment
  2. Fusion lattice computation
     - Selection of basis tree
     - Alternative verbalizations
     - Addition of frequent subtrees
     - Pruning of unique subtrees
  3. Linearization with LM
- Comparison on DUC data:
  - Outperforms baselines but beaten by humans on F-measure
  - Errors in grammaticality due to LM; good sentences scored badly



Sentence Fusion



As a standalone task

Sentence Fusion

As a standalone task

Marsi & Krahmer (2005)

- First introduced idea of *union* and *intersection* fusion
- Manual alignment of dependency trees used to compare against automatic alignment of Barzilay & McKeown (2005)
    - Very small dataset: one chapter of parallel Dutch translations of *Le Petit Prince*
    - Automatic alignment has 2x the recall of string-matching baseline
- Aligned node pairs manually labeled for *restatement*, *generalization* and *specification*
    - 50% of restatements and generalizations judged perfect
- LM-based ranking found to be of poor quality

E. Marsi & E. Krahmer (2005) Explorations in Sentence Fusion

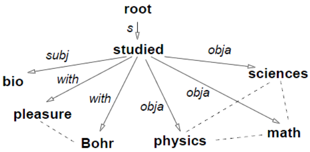
R. Barzilay & K. McKeown (2005) Sentence Fusion for Multidocument News Summarization

Sentence Fusion

As a standalone task

Filippova & Strube (2008)

- ILP over dependency graphs
  1. Simple alignment of dependency trees to create DAG
  2. Pruning using syntactic importance and word informativeness
  3. Globally optimal tree found with constraints for well-formedness
- Semantic constraints over:
  - Hyponyms and meronyms using GermaNet, taxonomy
  - Unrelated words through WikiRelate

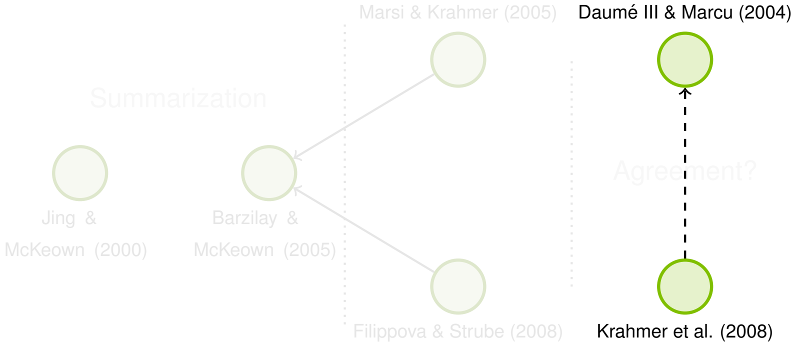


- Manual evaluation over German comparable biographies:
  - Significant readability gain over Barzilay & McKeown (2005); insignificant gain in informativity
  - Heavy reliance on parsing quality for good output

K. Filippova & M. Strube (2008) Sentence Fusion via Dependency Graph Compression

R. Barzilay & K. McKeown (2005) Sentence Fusion for Multidocument News Summarization

Sentence Fusion



Human agreement analysis

Sentence Fusion

Human agreement analysis

Daumé III & Marcu (2004)

- 50 sentence pairs each aligned to a single abstract sentence, compared with 3 human fusions and 3 baselines: longer sentence, truncated concatenation, Daumé III & Marcu (2002)
- No agreement of humans with reference under a *factoid* retrieval approach
  - Lack of document context?
- Longer baseline preferred to Daumé III & Marcu (2002) in absolute/relative ranking; vice versa for factoids
  - Grammaticality?
- Longer baseline outperforms reference in relative ranking
  - Out of context? Also, reference created by trained summarizer

H. Daumé III & D. Marcu (2004) Generic Sentence Fusion Is An Ill-Defined Summarization Task

H. Daumé III & D. Marcu (2002) A Noisy-Channel Model for Document Compression

Sentence Fusion

Human agreement analysis

Krahmer et al. (2008)

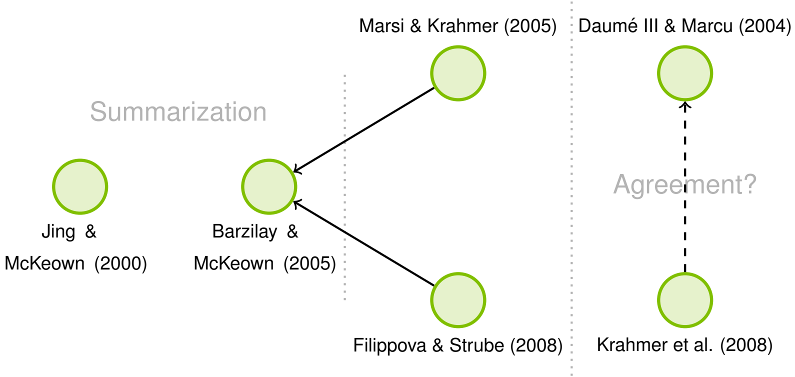
Dimensions: **query-based** vs **generic** fusion

intersection vs **union** fusion

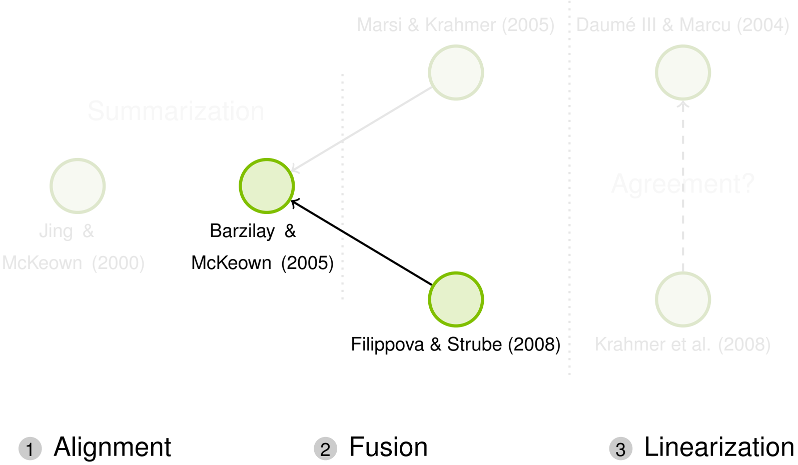
1. Data collection experiments over 25 medical questions:
   - Q-based fusion has less variation than generic fusion (by lower standard deviations over sentence length; more duplicates)
   - Q-based fusions do consistently better under Rouge metrics
2. Evaluation of preference on fusions of 20 questions:
   - Results show Q-based union $>$ Q-based intersection $>$ generic intersection/union
   - No significant difference between generic fusion types

E. Krahmer, E. Marsi & P. van Pelt (2008) Query-based Sentence Fusion is Better Defined and Leads to More Preferred Results than Generic Sentence Fusion

Sentence Fusion



Sentence Fusion



Alignment

Words:	Barzilay & Lee (2003)	Filippova & Strube (2008)	
Phrases (GIZA++):	Quirk et al. (2004)	Bannard & Callison-Burch (2005)	Callison-Burch (2008)
Syntactic parse:	Pang et al. (2003)		
Dependency parse:	Barzilay & McKeown (2005)	Marsi & Krahmer (2005)	

Alignment

Words:

Barzilay & Lee (2003) Filippova & Strube (2008)

Phrases (GIZA++):

Quirk et al. (2004) Bannard & Callison-Burch (2005) Callison-Burch (2008)

Syntactic parse:

Pang et al. (2003)

Dependency parse:

Barzilay & McKeown (2005) Marsi & Krahmer (2005)

Alignment

Words:

Barzilay & Lee (2003)

Filippova & Strube (2008)

Phrases (GIZA++):

Quirk et al. (2004)

Bannard & Callison-Burch (2005)

Callison-Burch (2008)

Syntactic parse:

Pang et al. (2003)

Dependency parse:

Barzilay & McKeown (2005)

Marsi & Krahmer (2005)

Alignment

Words:

- Barzilay & Lee (2003)
- Filippova & Strube (2008)

Phrases (GIZA++):

- Quirk et al. (2004)
- Bannard & Callison-Burch (2005)
- Callison-Burch (2008)

Syntactic parse:

- Pang et al. (2003)

Dependency parse:

- Barzilay & McKeown (2005)
- Marsi & Krahmer (2005)

Alignment

Words:

Barzilay & Lee (2003)

Filippova & Strube (2008)

Phrases (GIZA++):

Quirk et al. (2004)

Bannard & Callison-Burch (2005)

Callison-Burch (2008)

Syntactic parse:

Pang et al. (2003)

Dependency parse:

Barzilay & McKeown (2005)

Marsi & Krahmer (2005)

Alignment

Monolingual phrases

MacCartney et al. (2008)

- Exploits monolingual lexical and phrase-based resources for alignment
- MANLI: Phrase-based alignment system
  - Trained on annotated RTE2 alignments with averaged perceptron
  - Decodes alignments using a simulated annealing search for phrase segmentation
  - Uses lexical/phrasal similarity features and contextual features for position
- Human evaluation on RTE2 corpus:
  - Compares favorably to a bag-of-words aligner, GIZA++ and cross-EM MT aligners and the Stanford token-based aligner
  - Generates 21% of gold alignments exactly
  - Error analysis shows poor scoring of aligned phrases and NEs

---

B. MacCartney, M. Galley & C. Manning (2008) A Phrase-Based Alignment Model for Natural Language Inference

Linearization

LM-based:

Barzilay & McKeown (2005)

Marsi & Krahmer (2005)

Cohn & Lapata (2008)

Linearization

LM-based:

Barzilay & McKeown (2005)

Marsi & Krahmer (2005)

Cohn & Lapata (2008)

Linearization

Beyond language models

Filippova & Strube (2009)

- Proposes a distinction between phrase level and clause level
  1. Trigram LM to order within phrases
  2. Classifiers to order clause constituents
     - To determine best starting point of a sentence
     - To determine if precedence holds between adjacent constituents
- Evaluation on regenerating sentences from dependency trees:
  - Trigram LM has 76% accuracy at the phrase level but only 49% at the clause level
  - The combined approach has 67% accuracy at the clause level

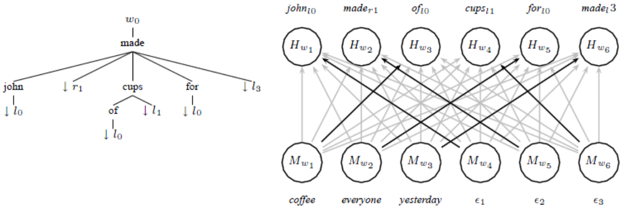
K. Filippova & M. Strube (2009) Tree Linearization in English: Improving Language Model Based Approaches

Linearization

Beyond language models

Wan et al. (2009)

- Linearization of dependency graph using minimum spanning tree (MST) formulation
    - Introduces direction labels in dependency graph; helpful for linearization
    - Approach based on assignment problem for argument satisfaction



- BLEU evaluation on string regeneration over PTB, BLLIP:
    - Argument-based approach beats vanilla MST algorithm by 7 points
    - Noun phrases not regenerated in task; this contributes 8-10 points

---

S. Wan, M. Dras, R. Dale & C. Paris (2009) Improving Grammaticality in Statistical Sentence Generation: Introducing a Dependency Spanning Tree Algorithm with an Argument Satisfaction Model

Conclusion

Text to text generation:

- Four primary tasks
  1. Paraphrasing
  2. Simplification
  3. Compression
  4. Fusion
- Various components
  1. Alignment
  2. Linearization
  3. Discourse
- Useful applications:
  1. Summarization
  2. Question answering
- Lots of techniques ...