Learning Sentential Paraphrases from Bilingual Parallel Corpora for Text-to-Text Generation

Juri Ganitkevitch, Chris Callison-Burch, Courtney Napoles, and Benjamin Van Durme

human language technology center of excellence

JOHNS HOPKINS UNIVERSITY

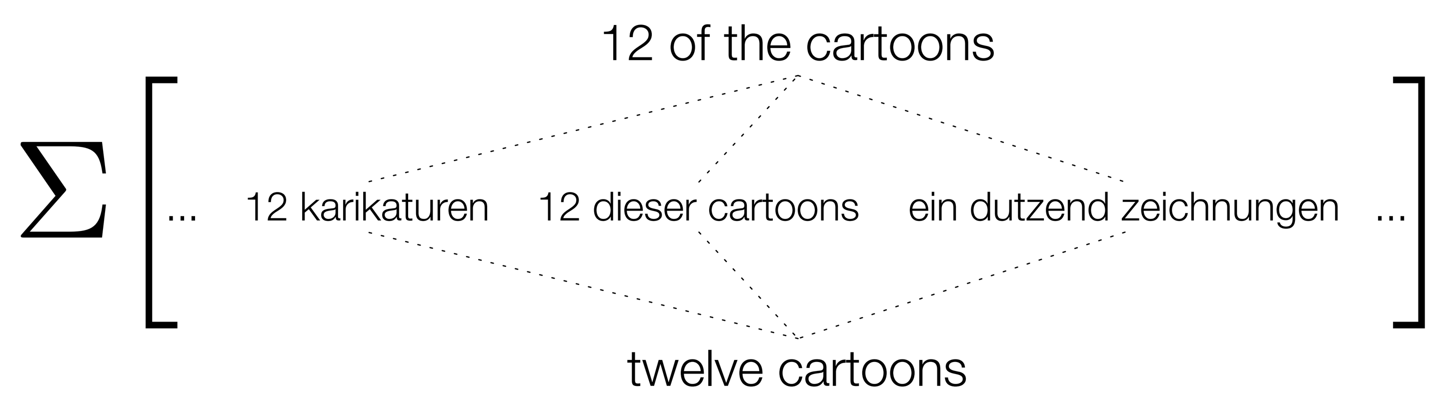
CENTER FOR LANGUAGE & SPEECH PROCESSING

Abstract

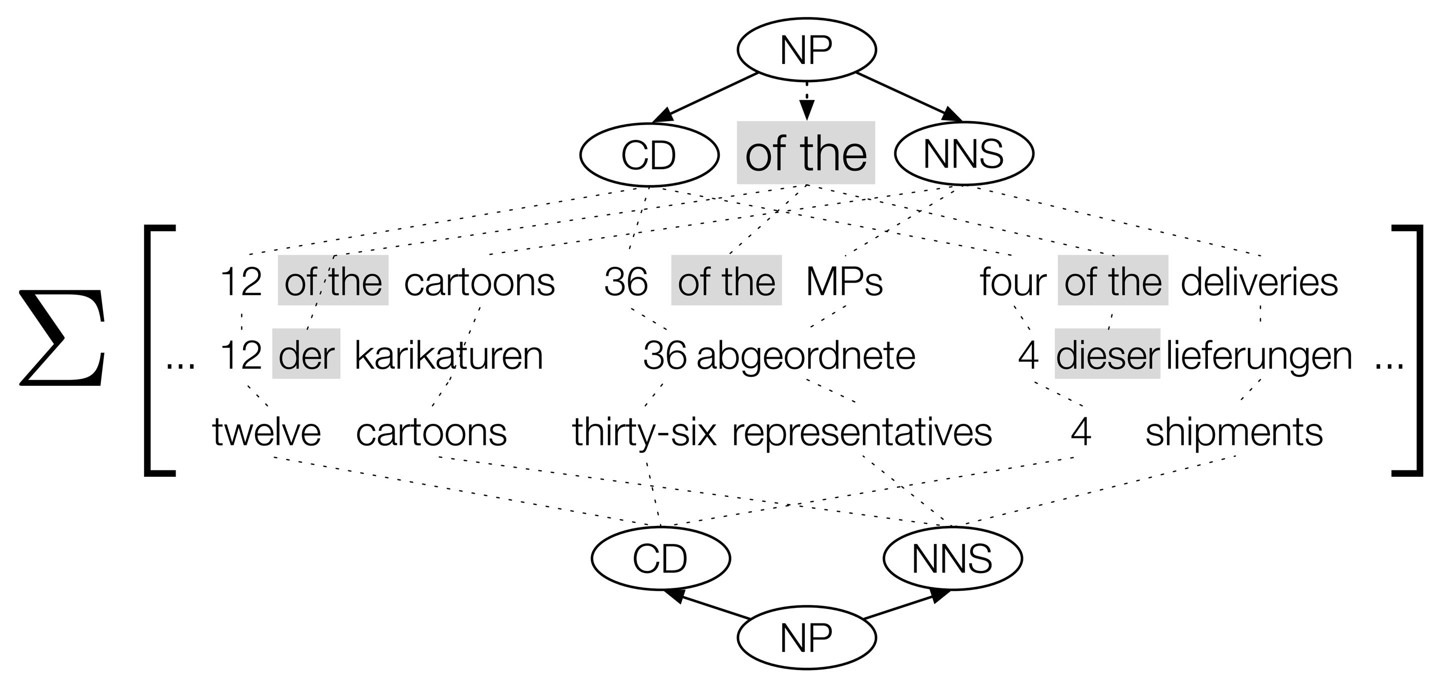
Previous work sucessfully extracted high quality phrasal paraphrases from bilingual parallel corpora. However, it is not clear whether bitexts can yield more sophisticated sentential paraphrases, that are more obviously learnable from monolingual parallel corpora. We extend bilingual paraphrase extraction to syntactic paraphrases and so are able to learn a variety of general paraphrastic transformations, such as passivization and dative shift. We discuss adapting our model to many text-to-text generation tasks by augmenting its feature set, development data, and parameter estimation routine. We illustrate this adaptation by using our paraphrase model for sentence compression and achieve results competitive with state-of-the-art compression systems.

Syntactic Paraphrases from Bitexts

When extracting phrasal paraphrases from a bitext, we pivot over the foreign sides in a translation phrase table and then aggregate probabilities over all common foreign phrases:



For syntactic paraphrases, we first extract syntactic translation SCFGs (i.e. rules with two right-hand sides and exact correspondence between the NTs on the right-hand side: "NP → CD of the NNS | CD dieser NNS"). We then analogously pivot and aggregate over the foreign side:



	Adapting from SMT..	..to Sentence Compression
Feature Functions	Phrasal and lexical probabilities quantify general paraphrase quality. More task-specific properties are not captured.	We add features that count the number of source and target words and the relative difference between them.
Dev Set	Tuning on English reference translations that are used to calculate BLEU for SMT. These are sentential paraphrases by definition, but do not reflect a particular task like compression.	We select pairs of sentences from a collection of multiple references that significantly differ in length. This allows us to obtain paraphrased compressions to use as development data.
Objective Function	Optimized for English-to-English BLEU score. The typically high inter-reference BLEU score causes the system to tune to self-paraphrasing.	We develop an objective function similar to BLEU, but with a "verbosity penalty" that allows a target compression rate to be set.
Augmentations	It is not typical for additional task-specific rules to be added in the standard SMT pipeline.	Additionally, we augment the grammar with deletion rules for specific POS (JJ, RB, DT) allowing for shorter quasi-paraphrastic compressions: JJ → superfluous | ε

Expressiveness of Paraphrases

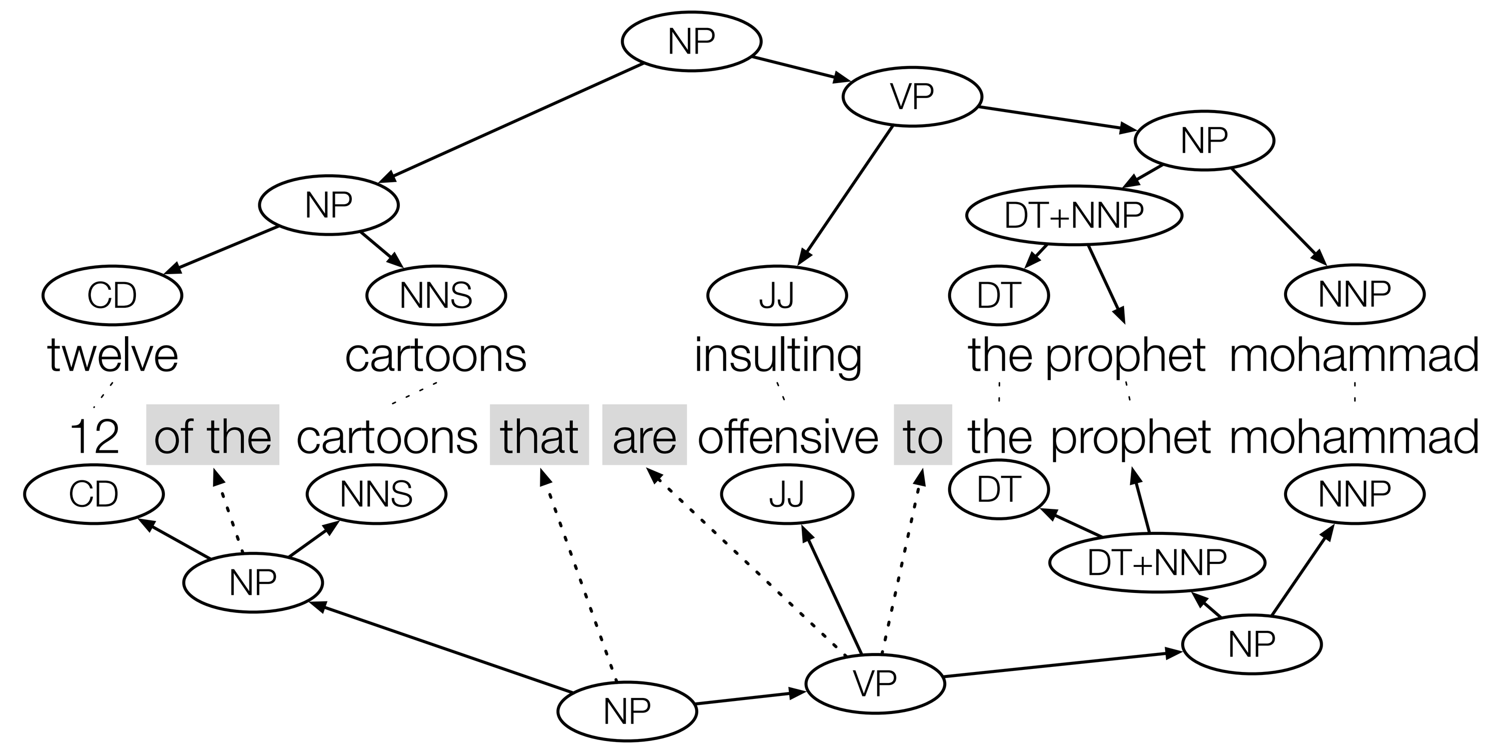
Our syntactic paraphrases capture a variety of meaning-preserving transforms:

Possessive rule	NP → NP →	the NN of the NNP | the NNP's NN the NNS_1 made by the NNS_2 | the NNS_2's NNS_1
Dative shift	VP → VP →	give NN to NP | give NP the NN provide NP_1 to NP_2 | give NP_2 NP_1
Adv./adj. phrase move	S/VP → S →	ADVP they VBP | they VBP ADVP it is ADJP VP | VP is ADJP
Verb particle shift	VP →	VB NP up | VB up NP
Reduced relative clause	SBAR/S → ADJP →	although PRP VBP that | although PRP VBP very JJ that S | JJ S
Partitive constructions	NP → NP →	CD of the NN | CD NN all DT\NP | all of the DT\NP
Topicalization	S →	NP, VP. | VP, NP.
Passivization	SBAR →	that NP had VBN | which was VBN by NP
Light verbs	VP → VP →	take action ADVP | to act ADVP to take a decision PP | to decide PP

Future Work

Our approach is highly flexible and can be extended to tasks such as sentence simplification, ESL error correction, legalese "translation", query expansion, question generation, RTE hypothesis generation and poetry generation.

Paraphrastic Sentence Compression



Paraphrase Rules

Lexical paraphrase:
JJ → offensive | insulting

Reduced relative clause:
NP → NP that VP | NP VP

Pred. adjective copula deletion:
VP → are JJ to NP | JJ NP

Partitive construction:
NP → CD of the NNS | CD NNS

Pivot Translation Rules

JJ → beleidigend | offensive
JJ → beleidigend | insulting

NP → NP die VP | NP VP
NP → NP die VP | NP that VP

VP → sind JJ für NP | are JJ to NP
VP → sind JJ für NP | JJ NP

NP → CD der NNS | CD of the NNS
NP → CD der NNS | CD NNS

Human Evaluation Results

We compare our system to state-of-the-art systems ILP (Clarke & Lapata, '08) and T3 (Cohn & Lapata, '07).