R3: Reverse, Retrieve, and Rank for Sarcasm Generation with Commonsense Knowledge

Tuhin Chakrabarty, Debanjan Ghosh, Smaranda Muresan and Nanyun Peng

Presented by Xiangjue Dong

August 6th, 2020

Characteristics of Generated Utterance

- Evaluative
- Reversal of valence between literal and intended meaning
- Semantic incongruity with context
- Target
- Relevant to communicative situation

Examples

Literal Input 1	I hate getting sick from fast food.
GenSarc 1	I love getting sick from fast food.
GenSarc 2	[I love getting sick from fast food.] [Stomach ache is just an additional side effect.]
Human 1	Shout out to the Mc donalds for giving me bad food and making me sick right before work in two hours.

Literal Input 2	I inherited unfavorable genes from my mother.
GenSarc 3	I inherited great genes from my mother.
GenSarc 4	[I inherited great genes from my mother.] [Ugly goes down to the bone.]
Human 2	Great I inherited all of my mother's GOOD genes.

Modules

- Reversal of valence
  - Negation
  - Lexical antonyms
  - E.g. I hate getting sick from fast food.
  - -> I love getting sick from fast food

Modules

- Reversal of valence
  - Negation
  - Lexical antonyms
- Retrieval of commonsense context
  - Generate relevant commonsense knowledge
  - Retrieve relevant sentences

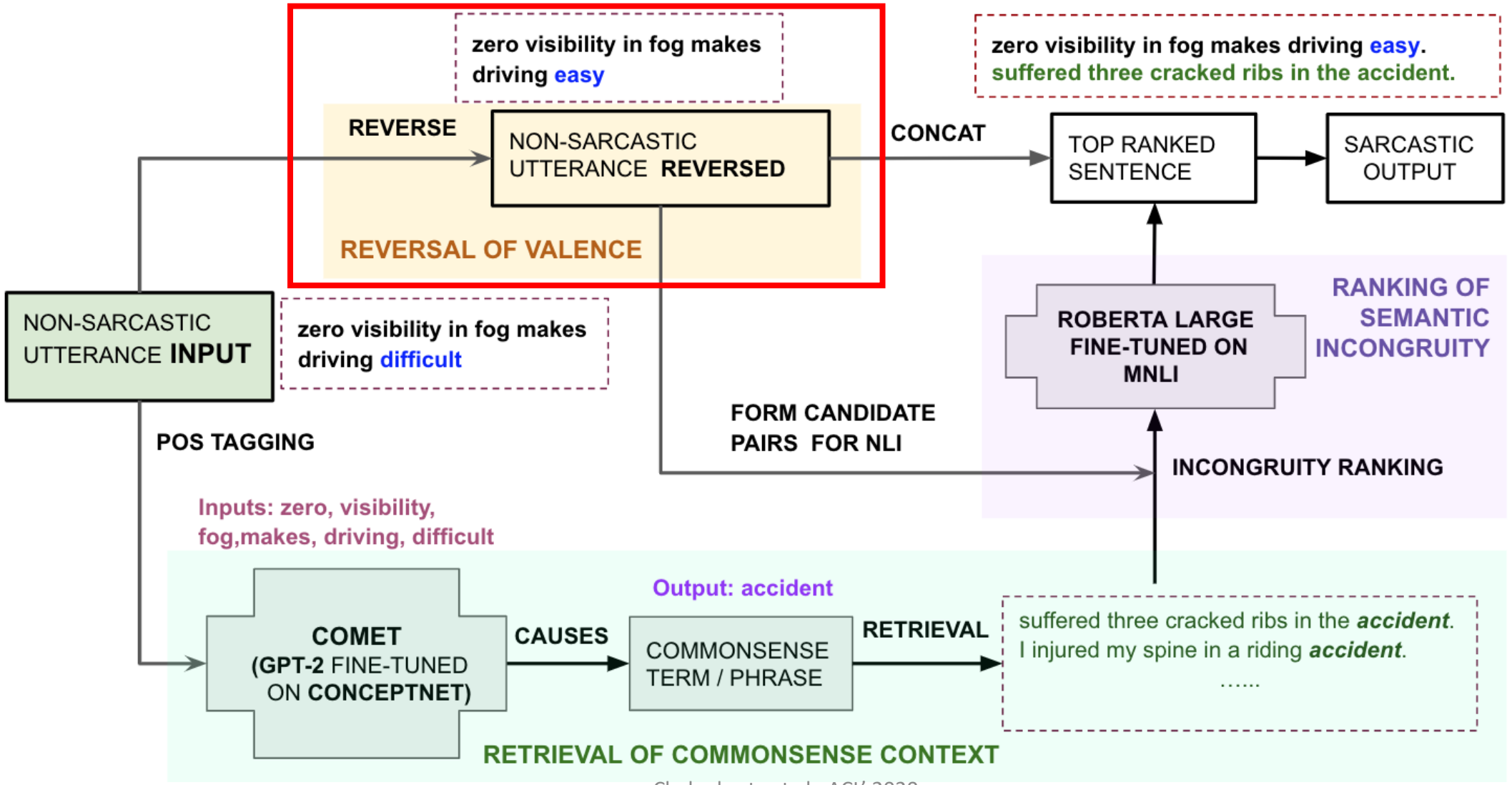
Modules

- Reversal of valence
  - Negation
  - Lexical antonyms
- Retrieval of commonsense context
  - Generate relevant commonsense knowledge
  - Retrieve relevant sentences
- Ranking of semantic incongruity
  - Measure contradiction between commonsense contexts and the sentence generated by the reversal of valence approach
  - Select the commonsense context that has the highest contradiction score

Sarcasm Factors Used in Generation

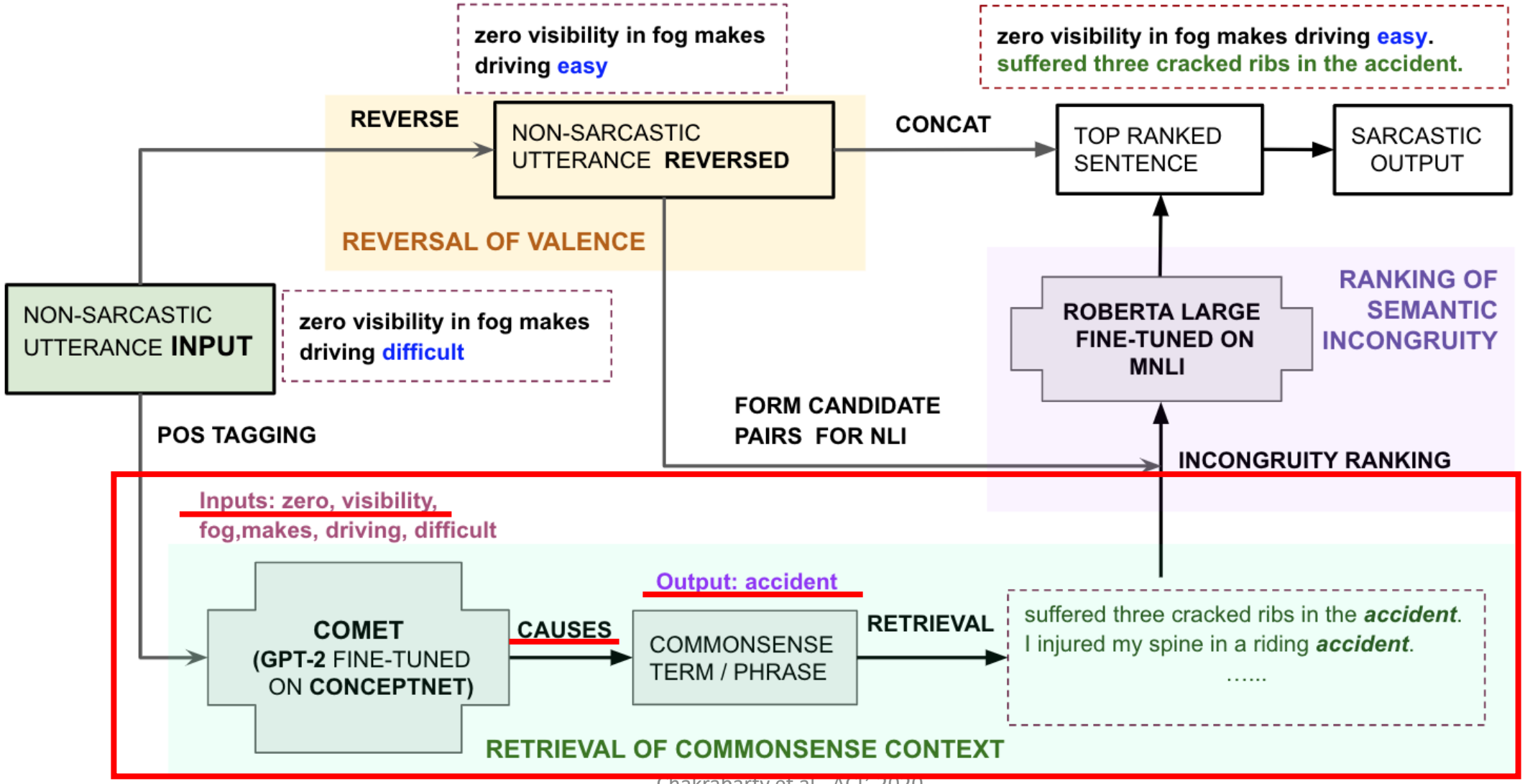
- Reversal of valence
  - Lexical antonym of negative sentiment words
  - Negation
- Semantic incongruity
  - E.g. “I love getting sick from fast food”.

Unsupervised Sarcasm Generation



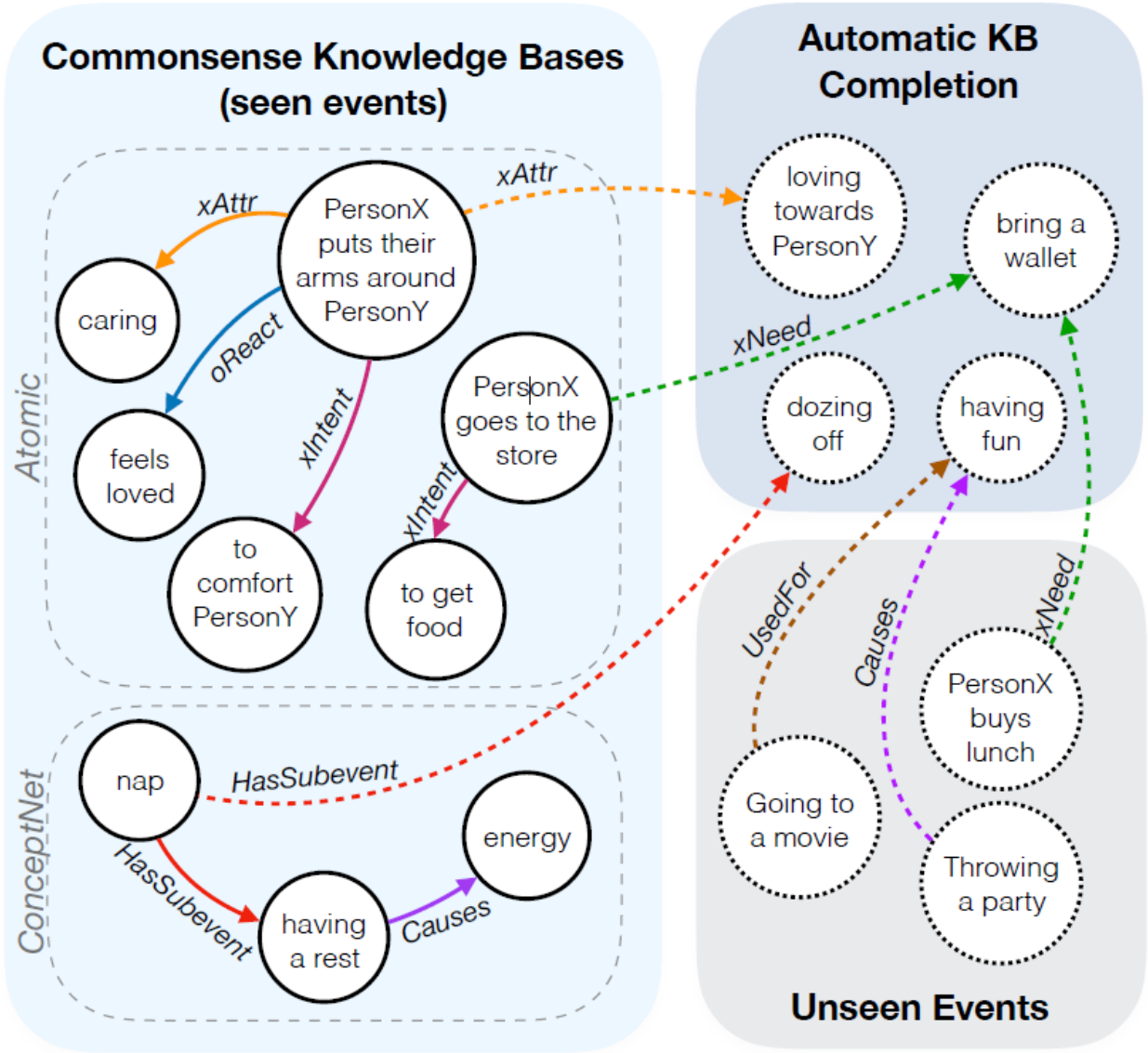
Chakrabarty et al., ACL' 2020

Unsupervised Sarcasm Generation



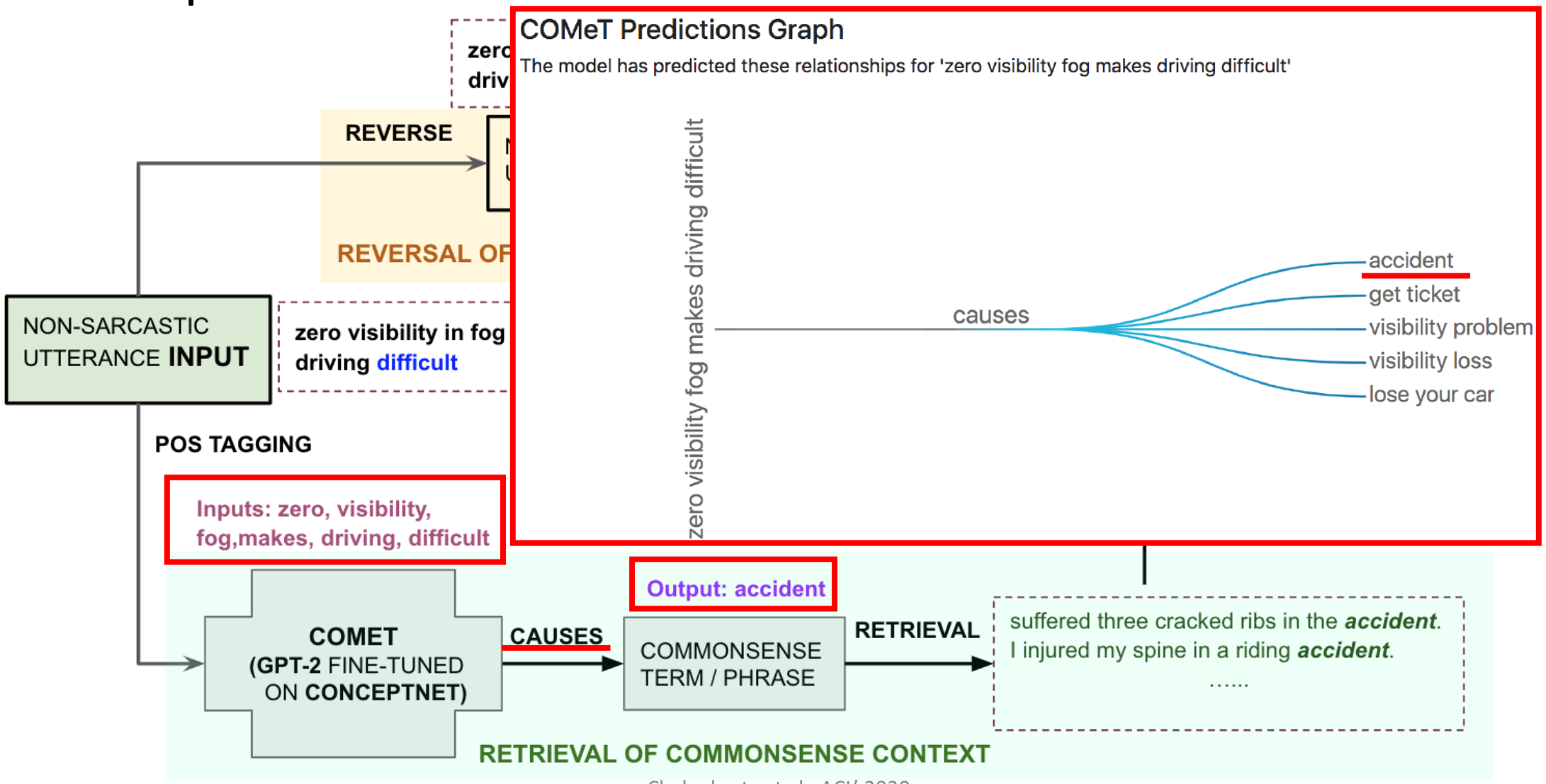
Chakrabarty et al., ACL' 2020

COMET



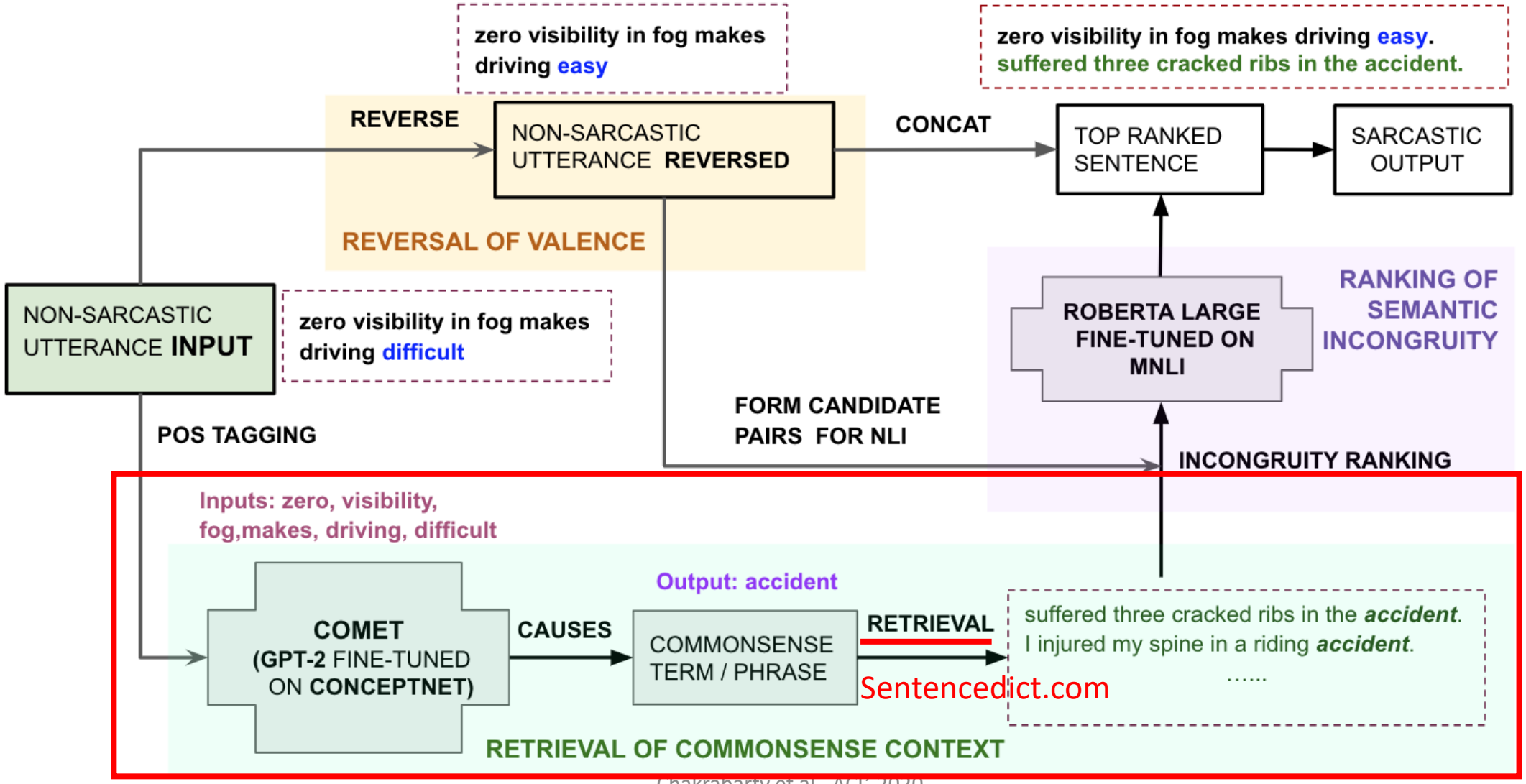
Bosselut et al., ACL' 2019

Unsupervised Sarcasm Generation



Chakrabarty et al., ACL' 2020

Unsupervised Sarcasm Generation

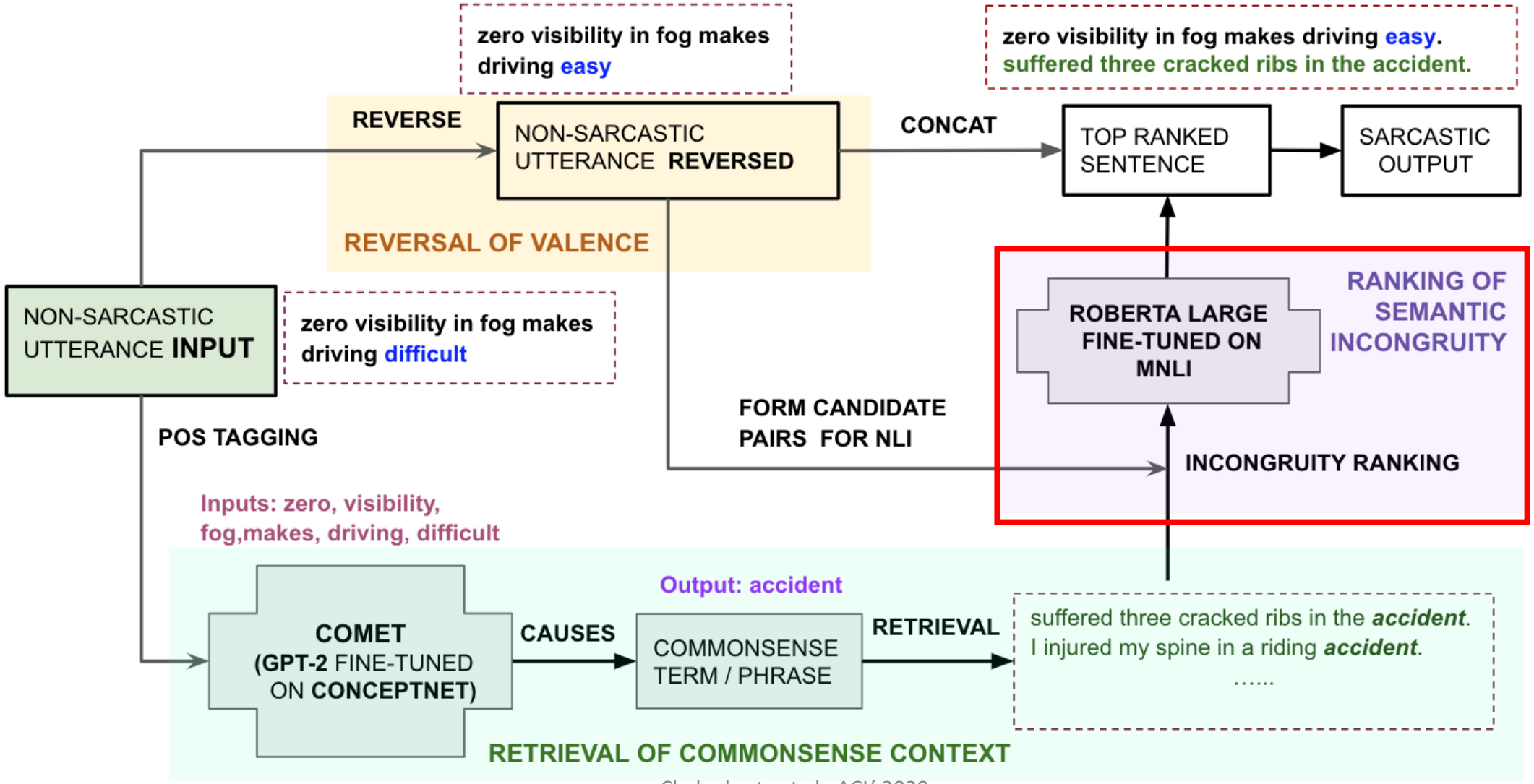


Chakrabarty et al., ACL' 2020

Grammatical Consistency

- Mismatched – modify
- Non-sarcastic input does not have any pronoun – change to “I”
  - E.g. input: Ignoring texts is literally the worst part of communication.
  - Retrieved: **He** has never suffered the torment of rejection.
  - Modified: **I** have never suffered the torment of rejection.
- Neural Grammatical Error Corrections System
  - Zhao, Wei et al. “Improving Grammatical Error Correction via Pre-Training a Copy-Augmented Architecture with Unlabeled Data.” *NAACL-HLT* (2019).

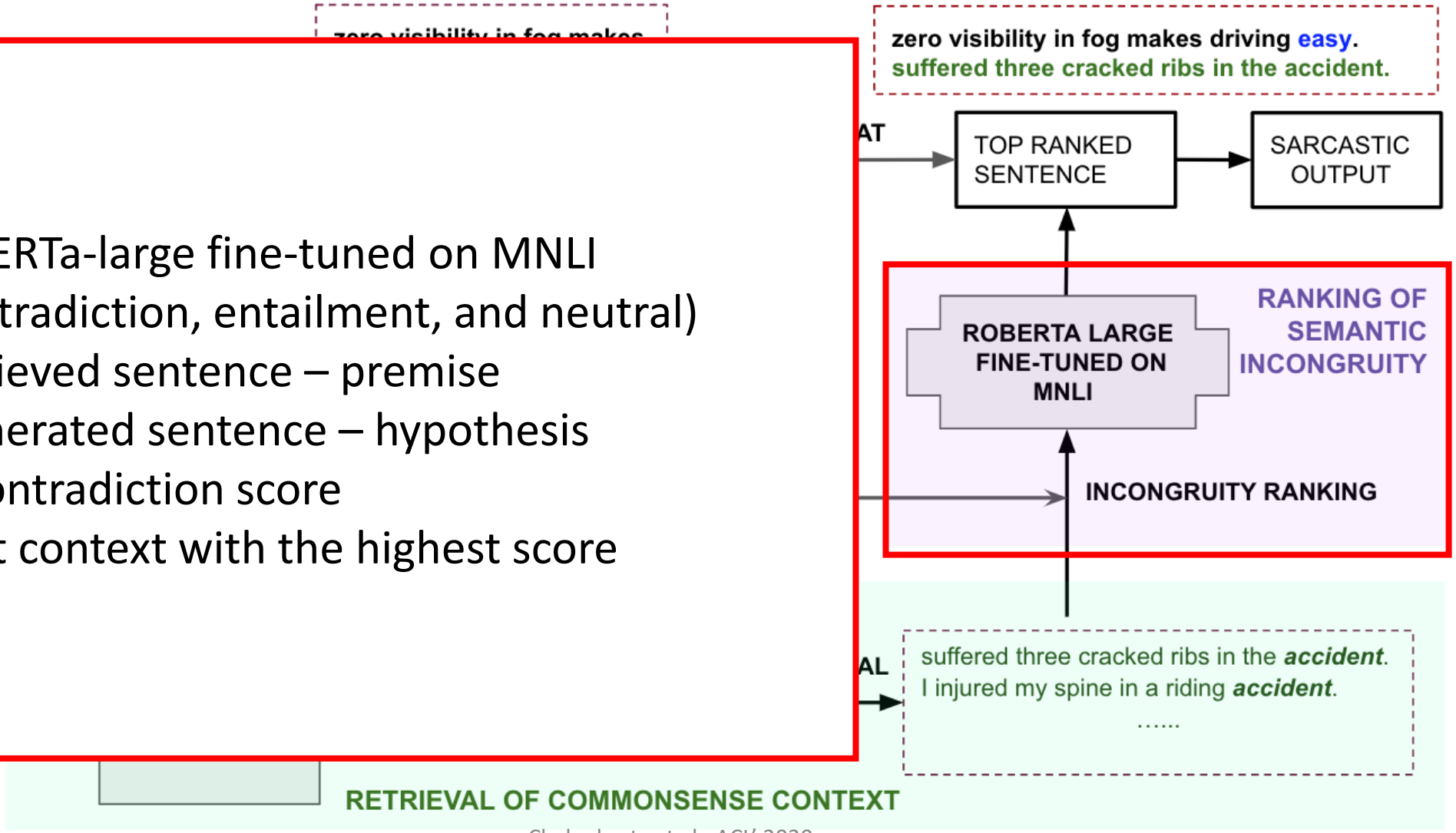
Unsupervised Sarcasm Generation



Chakrabarty et al., ACL' 2020

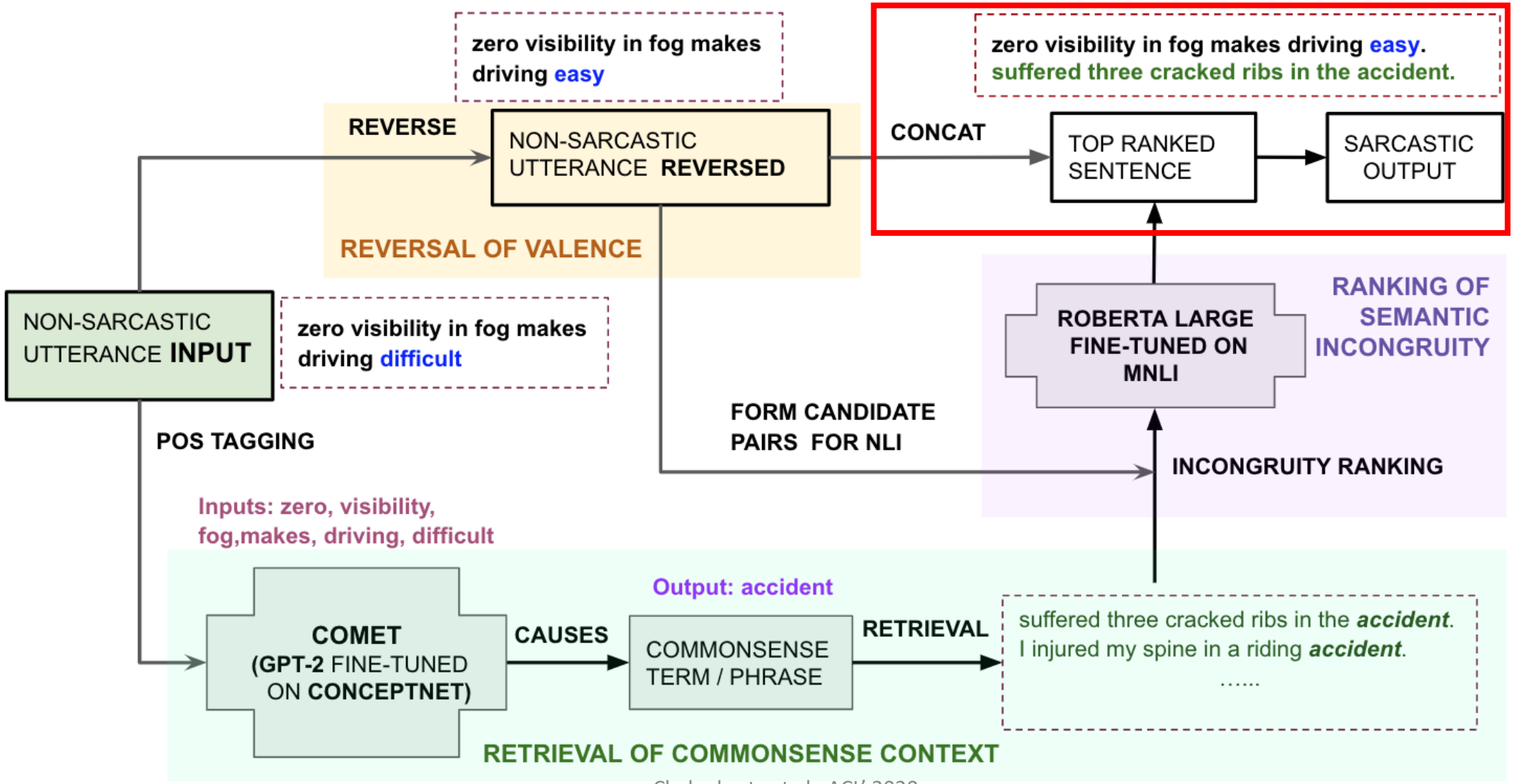
Unsupervised Sarcasm Generation

1) RoBERTa-large fine-tuned on MNLI (contradiction, entailment, and neutral)
2) Retrieved sentence – premise
   Generated sentence – hypothesis
   – contradiction score
3) Select context with the highest score



Chakrabarty et al., ACL' 2020

Unsupervised Sarcasm Generation



Chakrabarty et al., ACL' 2020

Experiments - Datasets

- merge Ghosh et al. (2020): 4762
- Peled and Reichart (2017): 3000
- non-sarcastic utterances: no longer than 15 words
- test set: 150

Experiments - Systems

- Full Model (FM)
- Reversal of Valence (RV)
- No Reversal of Valence (NoRV)
- No Semantic Incongruity (NSI)
- MTS2019: Mishra et al. (2019)
- Human (Gold) Sarcasm

Experiments – Evaluation Criteria

- Creativity
- Sarcasticness
- Humour
- Grammaticality

Experiments – Results

System	Sarcasticness	Creativity	Humor	Grammaticality
State-of-the-art (Mishra et al., 2019)	1.63	1.60	1.50	1.46
Human Generated	**3.57**	3.16	**3.18**	3.98
Reversal of Valence (RV)	3.00	2.80	2.72	**4.29**
No Reversal of Valence (NoRV)	1.79	2.28	2.09	3.91
No Semantic Incongruity (NSI)	3.04	2.99	2.90	3.68
Full Model (FM)	3.23*	**3.24**	3.08*	3.69

Experiments – Results

Aspect	FM vs Human		FM vs MTS2019	
	win%	lose%	win%	lose%
Sarcasticness	34.0	**55.3**	**90.0**	6.0
Creativity	**48.0**	36.0	**95.3**	4.0
Humor	40.6	**48.0**	**90.0**	4.0
Grammaticality	26.6	**56.6**	**98.0**	1.3

Future Work

- Build a model that can decide whether just the RV strategy is sufficient or if we need to add additional commonsense context to it

References

- Chakrabarty, T., Ghosh, D., Muresan, S., & Peng, N. (2020). R^3: Reverse, Retrieve, and Rank for Sarcasm Generation with Commonsense Knowledge. *ACL*.
- Bosselut, A., Rashkin, H., Sap, M., Malaviya, C., Çelikyilmaz, A., & Choi, Y. (2019). COMET: Commonsense Transformers for Automatic Knowledge Graph Construction. *ACL*.
- Zhao, W., Wang, L., Shen, K., Jia, R., & Liu, J. (2019). Improving Grammatical Error Correction via Pre-Training a Copy-Augmented Architecture with Unlabeled Data. *NAACL-HLT*.