NUS National University of Singapore | Computing

WING Web Information Retrieval / Natural Language Processing Group

The ELCo Dataset: Bridging Emoji and Lexical Composition

😄😂😇 🌉 abc ab cd AB CD

Zi Yun Yang*, Ziqing Zhang*, Yisong Miao*

http://github.com/WING-NUS/ELCo

Dataset and code @ Github

*All authors thank Prof Min-Yen Kan for his advising.

Victor Li 5:00 PM

🎉🐲 Happy Lunar New Year of the Dragon! 🐲🎉

Dear Friends, Group Mates and WING Alumni **@everyone**,

- Wishing all the group mates and alumni a fantastic start to the new year! This year, let the Dragon's legendary strength, wisdom, and luck bring us luck, 🍀 success in our studies 📚, achievements in our careers 🚀, and joy in every moment🥳.
- Let's make this year our best one yet! Go team! 🌟👩‍🎓👨‍🎓

🤔 Why are they written in this way?

🤔 What do they mean?

Excerpted from our group's Slack channel, with Victor's consent.

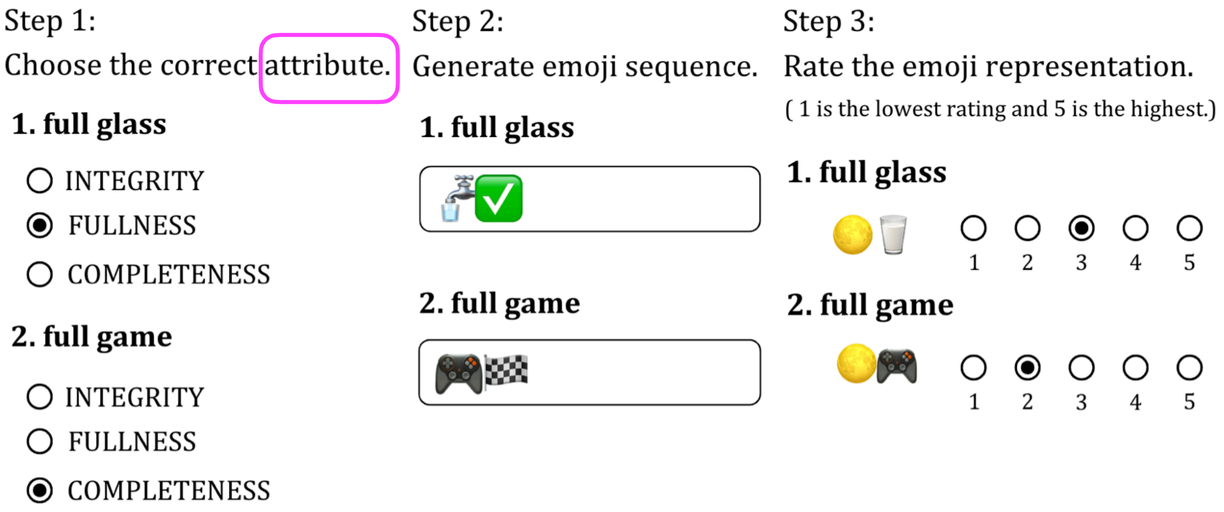
New Problem: Emoji Composition

In this paper, we address:

- **Dataset;**
- **Structure;**
- **Meaning;**

ELCo Dataset Creation

Annotation



Annotation workflow: ELCo's annotation process consists of three steps: (1) select the attribute of the phrase, (2) execute the annotation, and (3) rate the output from a rule-based system, Emojinating.

- HeiPLAS dataset: We choose 209 AN compounds encompassing 45 adjectives and 77 attributes.
- 40 university students (IRB approved).
- 1,655 responses received.

ELCo Dataset Creation

Validation

English Phrase	Attribute	ELCo's Annotations (length = 2)	ELCo's Annotations (length > 2)	Average length of ELCo's Annotation	Emojinating's Output	Emojinating Rating
full attention	INTEGRITY	💯❗ and 🧐✍️	💁‍♂️👂🏢	2.43	🟡🔔	3.1
full glass	FULLNESS	🥛🈵 and 🚰✅	🥛➕🚰🚰🚰🚰🚰🈵	2.71	🟡🥛	3.0
full game	COMPLETENESS	🎮🏁 and 🎮💯	💯🆓🎮	2.29	🟡🎮	2.7
full auditorium	FULLNESS	🈵🏢	🏘️🏘️🏢 and 🏢👦👨👱‍♀️🈵	4.14	🟡🚴	2.0
full life	FULLNESS	😊❤️	❤️👪💼💰💰💰💯 and 😆🥳💏	4.00	🟡😔	1.1

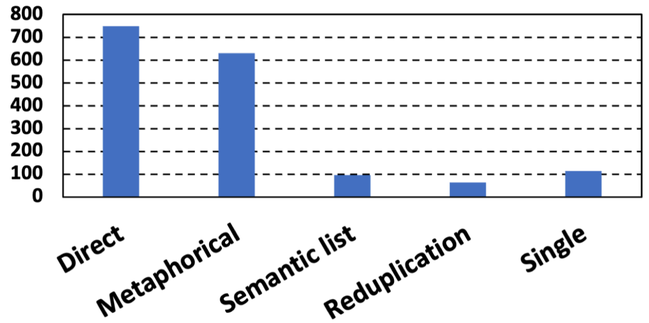
Compared with Emojinating:

- ELCo's **lengths?**
- Is ELCo being **literal?**
- Is ELCo more **metaphorical?**

Corpus study: Structures for Emoji Composition

	Compositional Strategy	EN Phrase	Emoji Sequence
Ex. 1	Direct	right man	✔️👨
Ex. 2	Metaphorical	right man	💏👫👦
Ex. 3	Metaphorical	clear explanation	👩💯🗣️
Ex. 4	Metaphorical	fresh bread	🎛️⏰🍞♨️
Ex. 5	Semantic list	bright future	😎👨👩👱
Ex. 6	Reduplication	big group	👫👫👫👫
Ex. 7	Single	right thing	✔️

Example for compositional structures.

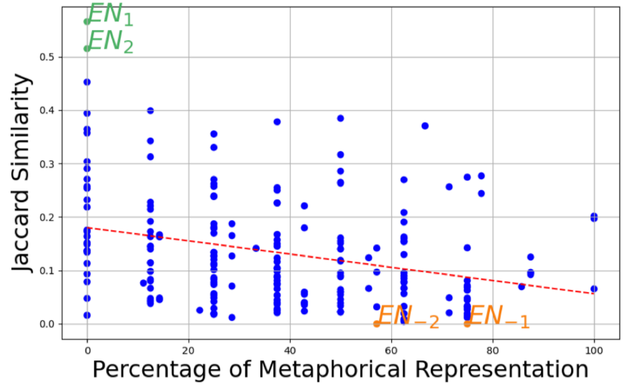


Corpus Study: Number of compositional structures identified in our corpus study (1,655 samples in total).

Inspired by Cohn's emoji grammar, following structures are identified:

- **Direct:** Translate;
- **Metaphorical:** Embody;
- **Semantic list:** Imply;
- **Reduplication:** Intensify;
- **Single:** Essence.

Corpus study: Structures for Emoji Composition

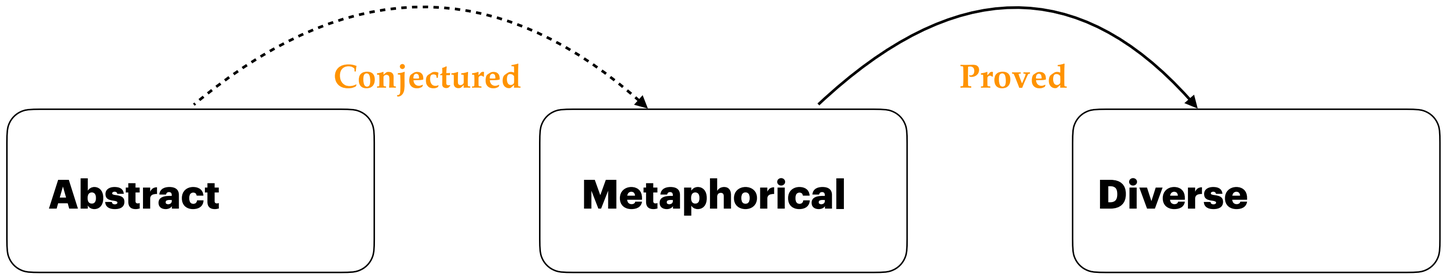


Jaccard Similarity: $J(A,B) = \frac{|A \cap B|}{|A \cup B|}$

- k = -0.123.

ID: EN PHRASE	Human Annotation Samples	Main Compositional Strategy	Jaccard Similarity
EN_1: WRONG MEDICINE	❌💊, 👩💊, ❌💊🤒, 👨💊	Direct	0.57
EN_2: WRONG ROAD	❌🛣️, 👩🛣️, 🤦🛣️, 👨🛣️	Direct	0.51
EN_{-2}: FAR SIDE	🧍🌎🧍, ↙️🛣️, 😪👉, 🚫➡️SOON	Metaphorical	0.0
EN_{-1}: IMMEDIATE INFLUENCE	👫↔️❗, 🐓🦹, 🙋🙋👱, ⏩👰👱	Metaphorical	0.0

More diverse.



Emoji-based Textual Entailment (EmoTE) Task

Formalization

Determines if a sequence of emojis *EM* (em_1 em_2 ... em_n) implies a English phrase *EN* (en_1 en_2 ... en_n). Formally:

- **Premise:** P em_1 em_2 ... em_n
- **Hypothesis:** P en_1 en_2 ... en_n

Input	Golden Label
Premise: This is 🧐✍️. Hypothesis: This is full attention.	Entailment
Premise: This is 🚰✅. Hypothesis: This is full attention.	Non-Entailment

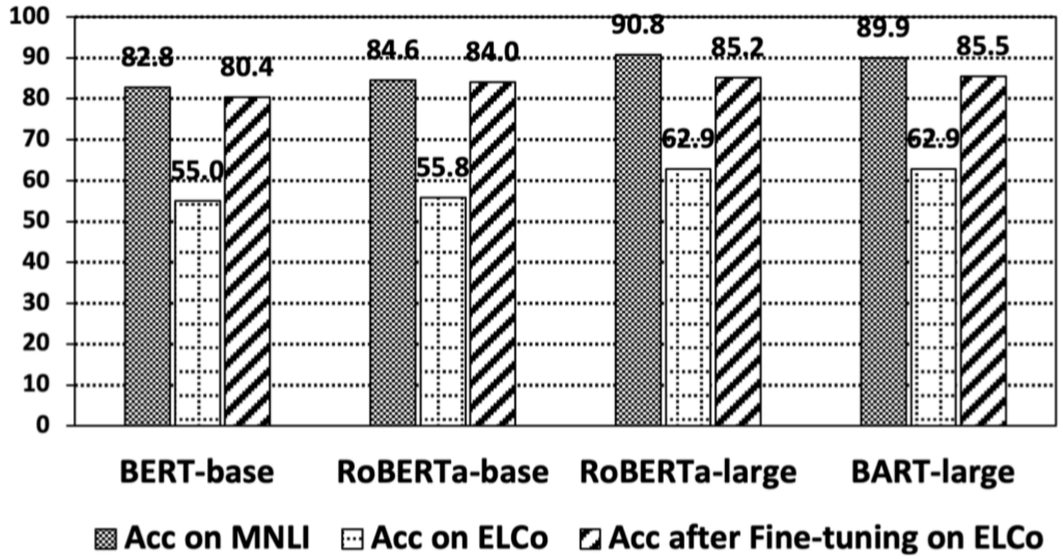
Emoji-based Textual Entailment (EmoTE) Task

Setup

- **Dataset split:** Roughly 70:15:15 for training, validation, and testing.
- **Sampling:**
  - Noun flipping by Shwartz and Dagan (2019): AN -> AN’ -> EM’
  - (AN, EM) and (AN, EM’)
- **Models:** BERT, RoBERTa, BART, ChatGPT-3.5.

Emoji-based Textual Entailment (EmoTE) Task

Overall performance



- Finding 1: Comparing with traditional NLI?
- Finding 2: Does fine-tuning help?

Emoji-based Textual Entailment (EmoTE) Task

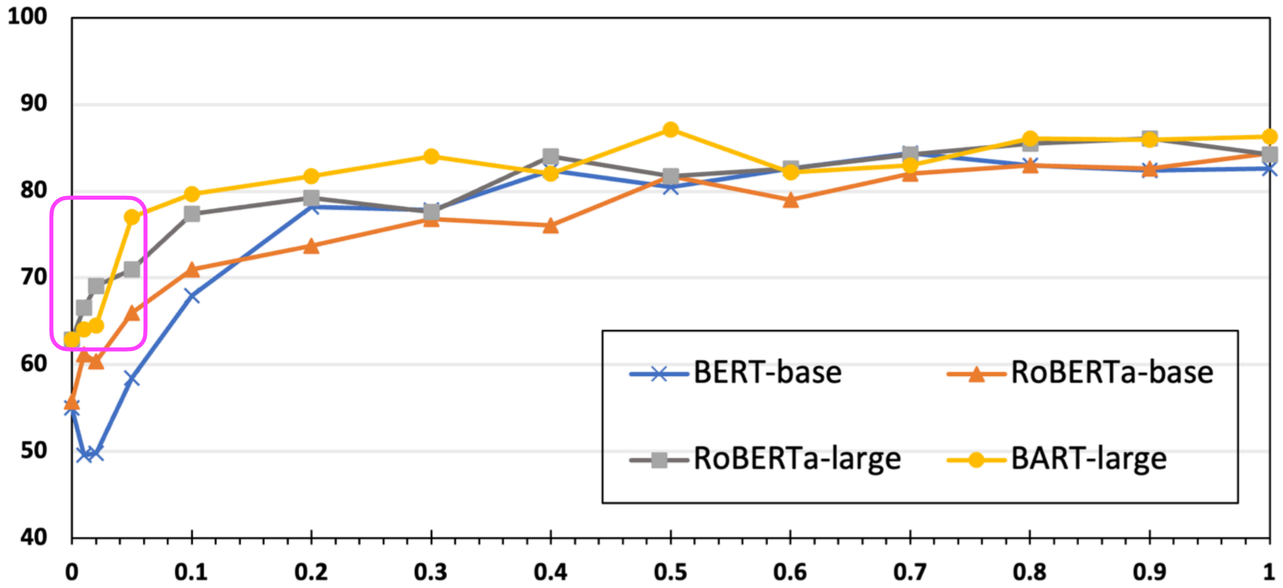
Fine-grained analysis

	w/o fine-tuning on ELCo					Fine-tuned on ELCo				
	$BERT_{base}$	$RoBERTa_{base}$	$RoBERTa_{large}$	$BART_{large}$	Avg	$BERT_{base}$	$RoBERTa_{base}$	$RoBERTa_{large}$	$BART_{large}$	Avg
Direct	34.7	35.6	41.5	51.7	40.9	$85.1_{2.3}$	$89_{3.3}$	$91.9_{2.2}$	$88.5_{2.8}$	88.6
Metaphorical	19.4	24.7	34.4	36.6	28.8	$68.4_{3.9}$	$73.1_{2.6}$	$80.4_{1.9}$	$82.8_{4.3}$	76.2
Semantic list	33.3	41.7	50.0	58.3	45.8	$86.7_{7.5}$	$78.3_{4.6}$	$85.0_{3.8}$	91.7_{0}	85.4
Reduplication	13.3	0	6.7	0	5.0	$65.4_{3.0}$	$52.0_{7.3}$	$62.7_{6.0}$	$88.0_{8.7}$	67.0
Single	19.0	19.0	19.0	52.4	27.4	$66.7_{3.4}$	$88.6_{2.6}$	$85.7_{4.8}$	$83.8_{2.6}$	81.2
Negative	83.4	83.0	90.3	82.2	84.7	$84.3_{4.1}$	$87.2_{1.9}$	$85.2_{0.7}$	$84.8_{1.2}$	85.4
Overall	55.0	55.8	62.9	62.9	59.2	$80.4_{1.5}$	$84.0_{0.8}$	$85.2_{0.9}$	$85.5_{0.9}$	83.8

- What's challenging?

Emoji-based Textual Entailment (EmoTE) Task

Scaling



- Can models converge?
- What models converge faster?

Emoji-based Textual Entailment (EmoTE) Task

Case study

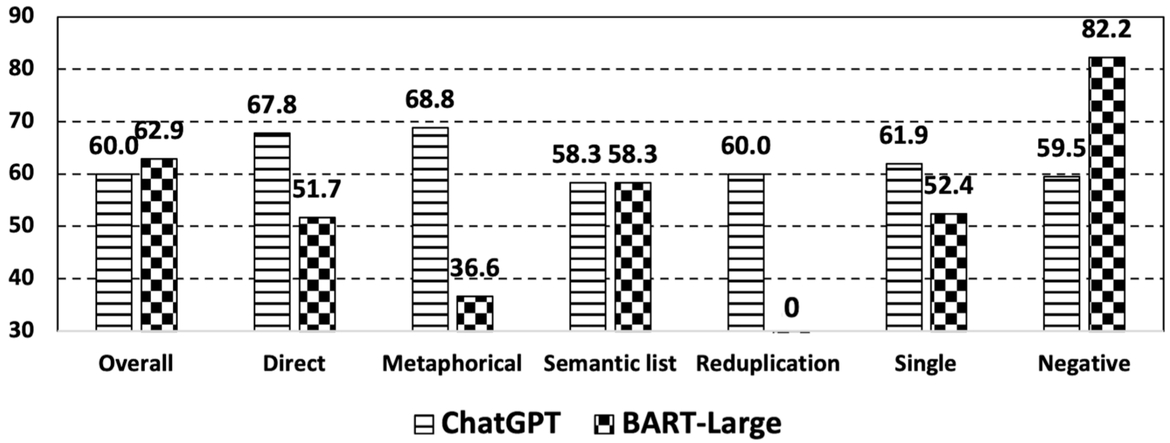
English	Emoji	Pre	Post	
1 big group		×	✓	
2 big city		×	✓	
3 hot forehead		×	✓	
4 thin soup		×	✓	
5 big city		×	×	→ Visual information of emoji.
6 ineffectual ruler		×	×	→ Commonsense knowledge.
7 full attention		×	×	→ Distant from EN phrase's meaning.
8 full life		×	×	→ Distant from EN phrase's meaning.

Emoji-based Textual Entailment (EmoTE) Task

ChatGPT

$BERT_{base}$	$RoBERTa_{base}$	$RoBERTa_{large}$	$BART_{large}$	ChatGPT
55.0	55.8	62.9↑	62.9↑	60.0

Overall Performance comparison of ChatGPT



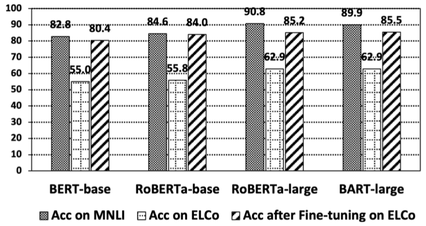
Fine-grained Performance comparison of ChatGPT

- Is ChatGPT competitive?
- What is ChatGPT good at?

Conclusion

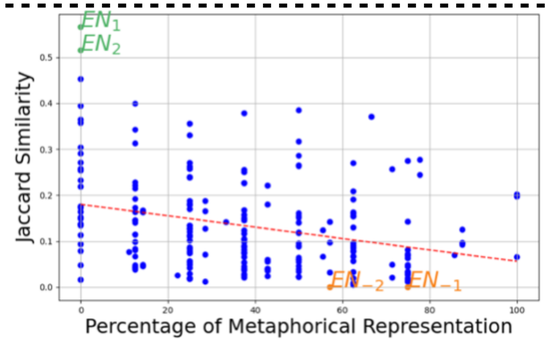
English Phrase	Attribute	ELCo's Annotations (length = 2)	ELCo's Annotations (length > 2)	Average length of ELCo's Annotation	Emojinating's Output	Emojinating Rating
full attention	INTEGRITY	💯❗ and 🧐✍️	🤷👂🏛️	2.43	🟡🔔	3.1
full glass	FULLNESS	🥛🈵 and 🍶✅	🥛➕🍶🍶🍶🍶🈵	2.71	🟡🥛	3.0
full game	COMPLETENESS	🎮🏁 and 🎮💯	💯🆓🎮	2.29	🟡🎮	2.7
full auditorium	FULLNESS	🈵🏛️	🏚️🏚️🏛️ and 💺😲😲😲👴🈵	4.14	🟡🎭	2.0
full life	FULLNESS	😊❤️	❤️👰🏠💰💰💰💯 and 😆🥳🍻	4.00	🟡😆	1.1

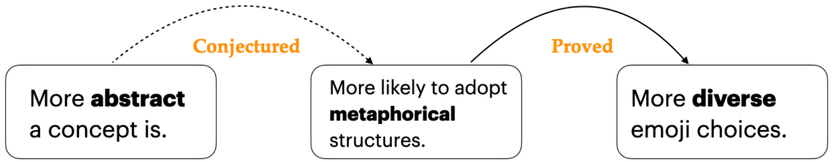
ELCo dataset is comprised of 1,655 annotations of 209 EN phrases 45 adjectives and 77 attributes.



English	Emoji	Pre	Post
1 big group	🤗🤗🤗🤗🤗	×	✓
2 big city	🌇🏙️🌇🏙️🌆🗽	×	✓
3 hot forehead	🥵😘👨	×	✓
4 thin soup	💦💦🍜	×	✓
5 big city	🌆🌆🌆🌆🌆	×	×
6 ineffectual ruler	🤷🤷🤷👎👎	×	×
7 full attention	🧐✍️	×	×
8 full life	🥰😆🏡	×	×

EmoTE is challenging for all models, but fine-tuning on ELCo helps to learn useful emoji composition skills.





Corpus study reveals five structures to compose emoji compositions, and we show metaphorical structures use more diverse emojis.

References

Matthias Hartung. 2015. *Distributional Semantic Models of Attribute Meaning in Adjectives and Nouns.* Ph.D. thesis.

João Miguel Cunha, Pedro Martins, and Penousal Machado. 2018. *Emojinating: Representing concepts using emoji.* In Workshop Proceedings from The 26th International Conference on Case-Based Reasoning (ICCBR 2018), Stockholm, Sweden, volume 185.

Qing Lyu, Zheng Hua, Daoxin Li, Li Zhang, Marianna Apidianaki, and Chris Callison-Burch. 2022. *Is "my favorite new movie" my favorite movie? probing the understanding of recursive noun phrases*. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 5286–5302.

Vered Shwartz and Ido Dagan. 2019. *Still a pain in the neck: Evaluating text representations on lexical composition.* Transactions of the Association for Computational Linguistics, 7:403–419.

Acknowledgements

The authors are grateful to their advisor, Prof. Min-Yen Kan, for his great support for this project. This includes his contributions to the initial ideation and project scoping, the technical insights he provided during our frequent meetings, his detailed feedback on our manuscript, and his unwavering support despite earlier rejections. Without him, this paper could not have been published.

The authors also extend their thanks to João Miguel Cunha and Qing Lyu for their technical consultation via emails about their related works. The authors also thank anonymous reviewers for their valuable comments.