Automatic Extraction of Relationships among Motivations, Emotions and Actions from Natural Language Texts

Fei Yang

Proceedings of the 38th Pacific Asia Conference on
Language, Information and Computation (PACLIC 38)

Nathaniel Oco, Shirley N. Dita, Ariane Macalinga Borlongan, Jong-Bok Kim (eds.)

2024

Automatic Extraction of Relationships among Motivations, Emotions and Actions from Natural Language Texts

Fei Yang
CogBeauty Lab, China
yftadyz@163.com

Abstract

We propose a new graph-based framework to reveal relationships among motivations, emotions and actions explicitly given natural language texts. A directed acyclic graph is designed to describe human's nature. Nurture beliefs are incorporated to connect outside events and the human's nature graph. No annotation resources are required due to the power of large language models. Totally 92,990 relationship graphs are extracted from food reviews, of which 63% make logical sense. We make further analysis to investigate error types for optimization direction in future research.

1 Introduction

In daily life, different motivations drive humans to produce different behaviors, and at the same time, the satisfaction of motivations leads to different emotions. Understanding relationships among motivations, emotions, and subsequent actions has drawn a lot of attentions in the research community. One prevailing practice is, given an event text, annotators generate description texts of its motivations, emotions, and subsequent actions (Rashkin et al., 2018; Sap et al., 2019; Ghosal et al., 2022). Or annotators mark motivations, emotions, and actions on its contexts (Poria et al., 2021; Mostafazadeh et al., 2020; Gui et al., 2018). Then deep learning models are trained over the generated or labeled datasets, which encode the relationships into the models' parameters. One drawback of this paradigm is, it fails to reveal relationships explicitly, providing not much help in understanding human intelligence, although these black-box models perform well in real applications. Another drawback is, it heavily relies on human resources and workflow designs for annotation.

In this work, we propose a framework to explicitly handle relationships among motivations, emotions and actions, which automatically generates directed acyclic graphs (MEA-DAG) given natural language texts. By drawing on findings from cognitive science, a Nature Design graph is built manually, which reveals human's inside nature, being formed through thousands of years of genetic evolution. Our framework also incorporates Nurture Belief, learned from developmental experiences. Nurture Belief plays a key role in connecting outside world events and Nature Design. Figure 1 shows the Nature Design graph and a MEA-DAG example. Large language model (LLM) is used to extract and improve the quality of Nurture Belief. Therefore no annotation resources are required in our framework, and efforts are put on prompt engineering instead of annotation workflow design.

To reduce the complexity of the problem, only the motivation of human's need for food (Maslow, 1943) is focused. We divide this motivation into two types: positive and negative, which correspond to food need being met and not met respectively. From review texts of Amazon Fine Foods Reviews (McAuley and Leskovec, 2013), totally 92,990 MEA-DAGs are extracted out and 63% of them make logical sense. Error analysis is implemented to investigate the error types and find future research directions. All codes and data are released publicly. [1]

2 Related Work

Event2Mind (Rashkin et al., 2018) asks annotators to provide short textual descriptions of motivations and emotional reactions given an event. The collected texts serve as the training set of a encoder-decoder model, which predicts motivation and emotion over a new event. ATOMIC (Sap et al., 2019) extends the annotation dimensions, and trains a encoder-decoder model for inference. In (Ghosal et al., 2022), given an event in a dialogue, annotators answer five dimensions: cause, subsequent

[1] https://github.com/yftadyz0610/MAE-DAG

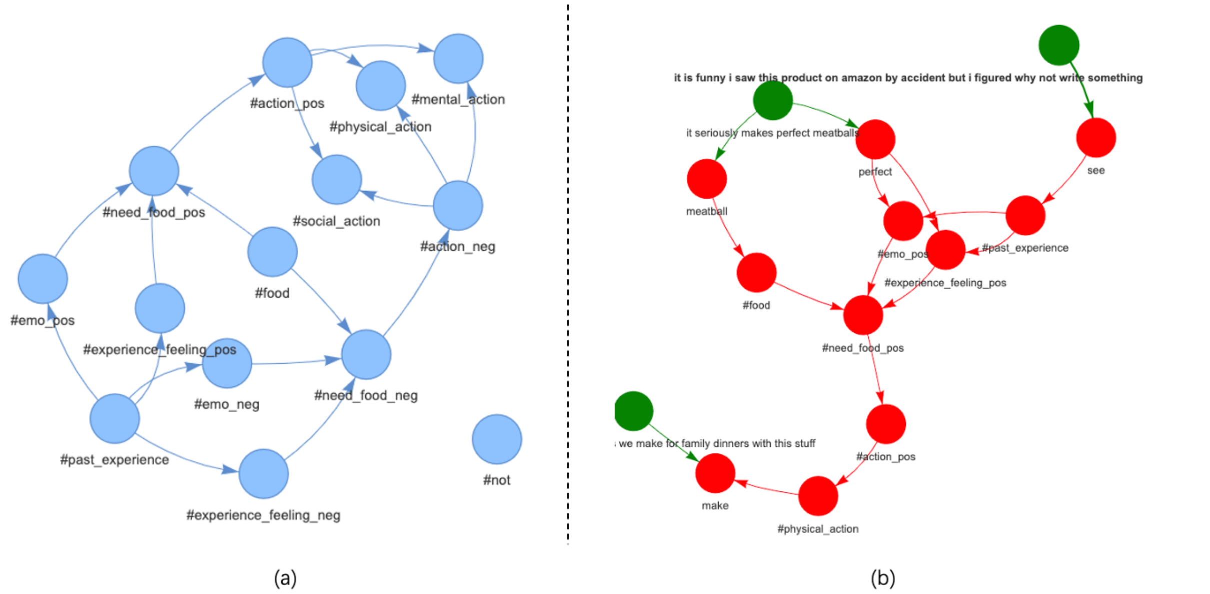


Figure 1: (a) Nature Design. This graph reveals the interactive mechanism among motivations, emotions and actions in human's nature. (b) An MEA-DAG example. Events are extracted by ASER from a review and presented in green color. The activated nodes of Nature Design are in red color. Other nodes are omitted. Nurture Belief participates in linking events to corresponding nodes, and presented in the MEA-DAG as well. For instance, "it seriously makes perfect meatballs", has a connection with `#food` by the belief tuple ("meatball", `#food`).

event, prerequisite, motivation, and emotional reaction for tuning transformer-based models. Instead of generating artificial answers, other researchers choose to mark key information directly on the contexts of events. Poria et al. (2021) annotates the cause of emotions manually at phrase level in two conversation datasets, and transformer models are fine-tuned for inference. Gui et al. (2018) annotates the cause of emotions from emotional context, and then a convolution kernel-based model is learned. In (Mostafazadeh et al., 2020), given a sentence and its context, ten dimensions including motivation, emotion, other implicit causes, and its effect are annotated by crowdsourcing. They then train a encoder-decoder model to infer both specific statements and general rules for new scenarios. All these methods code the relationship among motivations, emotions and actions into parameters of their models in a black-box style. Therefore, they contribute very little to the understanding of this relationship, although their models could do excellent inference on new scenarios.

3 Methods

Our framework consists of four phases: (1) loading Nature Design and Nurture Belief, (2) perceiving states, (3) forward transmitting, and (4) taking actions, which are shown in Figure 2. It mimics human brain's cognition process. First of all, a brain stores innate evolutional design and acquired developmental experiences. Next, suppose an event occurs around, the brain perceives this event, then neutrons send signals along axons and dendrites, and finally an action is taken by invoking body parts. Afterwards, the brain perceives feedback from the action, which forms a closed loop of cognition, providing abilities to (1) form new knowledge about this world, and (2) guide next action based on all existed knowledge in the brain. The feedback loop is not covered in this work and left for future research.

3.1 Loading Nature Design and Nurture Belief

Nature Design starts from `#past_experience`, whose behavioral outcomes drive emotions and feelings. Emotions and feelings are involuntary, which serve as passive states, reflecting patterns of physiological activities (Panksepp, 2004). After perceiving these passive states, we infer whether human's need of food is satisfied or not. Positive feelings or emotions mean the need is satisfied while negatives mean the opposite, which are shown as directed links in Figure 1 (a). Positive actions are driven to strengthen being able to continuously meet the need when it's satisfied. Negative ac-

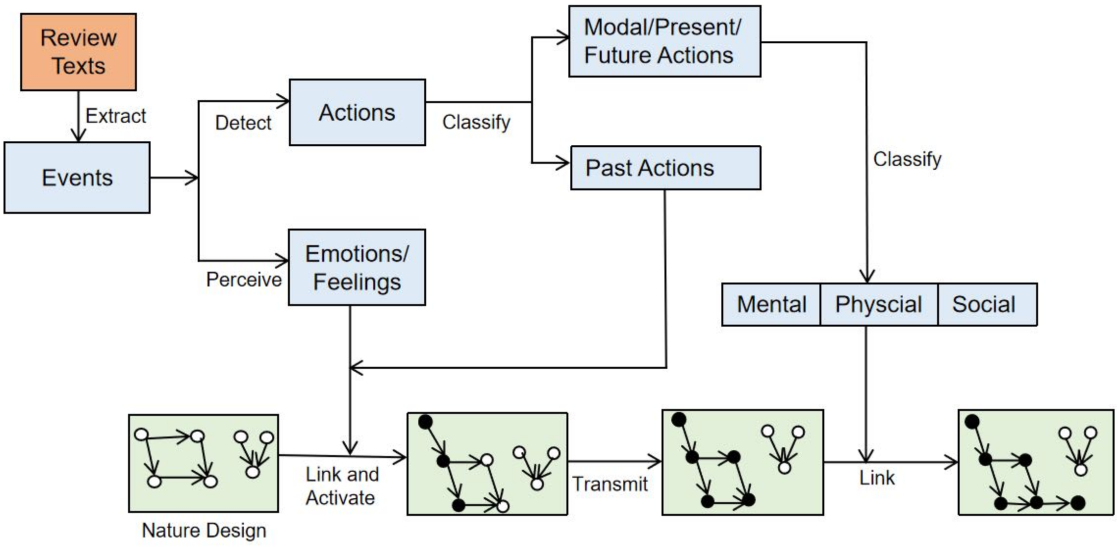


Figure 2: Framework of computing a MEA-DAG. It inputs the text of a food review (top left corner) and outputs a graph (bottom right corner). The green line at the bottom shows the evolution of a MEA-DAG in different processing stages, which imitates human brain's cognition process.

tions are driven to prevent it from happening again when human's need is dissatisfied. All actions are further broken down into three types: `Mental`, `Physical` and `Social`. Nodes of Nature Design are regarded as innate, different from learned experience (Izard, 1992; Deci and Ryan, 2000). Node definitions are summarized in Table 1. We admit that actions are not only determined by motivations, but also biologically, culturally, and situationally determined as well (Maslow, 1958). We ignore these factors and explore them in future research.

Nurture Belief includes three parts: food entities, experience feelings and emotions. They connect real world descriptions to abstractive concepts, stored as a set of tuples {("word", `node`)}. For instance, "cheerful" describes a positive emotion, which is stored as ("cheerful", `#emo_pos`). When an event is linked to a node, the related Nurture Belief tuples are presented in its MEA-DAG as well.

Food Entities. WordNet (Miller, 1995) provides an ability to link concrete entities to abstract categories. We start from the word "food" and find all its hyponyms. The hyponyms with food-unrelated senses are removed to improve accuracy. Totally 1,842 tuples of ("word", `#food`) are collected.

Experience Feelings. SentiWordNet (Baccianella et al., 2010) provides positive, negative and neutral feeling scores at sense level. By setting `PosScore`>0.6 and `NegScore`>0.6, positive and negative senses are extracted out respectively. Adjectives with only positive senses are classified as `#experience_feeling_pos`, and negative-senses only are classified as `#experience_feeling_neg`. GLM-4 (GLM et al., 2024), an open-source LLM,

- Emotion Classification Prompt

Judge if the INPUT word is describing positive/negative emotion. Answer in English. Explain your reasoning then state the answer. Return =True= or =False=. INPUT: word text OUTPUT:

- Negative Feeling Detection Prompt

Judge if the INPUT word is describing terrible experience. Answer in English. Explain your reasoning then state the answer. Return =True= or =False=. INPUT: word text OUTPUT:

- Action Type Classification Prompt

Mental: a mental action happens inside human beings and isn't visible. You engage in covert behaviour when you think since no one can see you thinking. Physical: a physical action is a behaviour that's visible and happens outside of human beings. Examples of overt behaviour include eating or drinking something and taking part in sports, such as football or riding a bicycle. Social: social behavior accounts for actions directed at others. It is concerned with the considerable influence of social interaction and culture, as well as ethics, interpersonal relationships, politics, and conflict. Judge the class that the INPUT action belongs to. Answer in English. Explain your reasoning then state the answer. Return =Mental= or =Physical= or =Social=. INPUT: word text OUTPUT:

Figure 3: Prompt engineering. We rely on LLM to accelerate the establishment of Nurture Belief, and no longer rely on manual labeling resources.

Table 1: Explanation of nodes in Nature Design.

Node	Explanation
#food	Food entities, e.g. bread, apple.
#experience_feeling_pos	Positive feelings, e.g. delicious, easy.
#experience_feeling_neg	Negative feelings, e.g. bitter, hard.
#emo_pos	Positive emotions, e.g. happy, cheerful.
#emo_neg	Negative emotions, e.g. sad, angry.
#need_food_pos	Human's need of food is satisfied.
#need_food_neg	Human's need of food is dissatisfied.
#past_experience	Actions that take place in the past and result in a change of need state, e.g. bought, searched. It's the root node of Nature Design.
#action_pos	Actions that are driven to strengthen being able to continuously meet the need.
#action_neg	Actions that are driven to prevent it from happening again when human's need is dissatisfied.
#mental_action	Actions that happen inside human beings, not visible, e.g. analyze, versify.
#physical_action	Actions that happen outside human beings and are visible, e.g. wash, peel.
#social_action	Actions that are directed at others, e.g. denounce, rent.

is used over the negative adjectives to filter out bad cases. We list the prompt in Figure 3. Totally 1,415 positive tuples and 1,239 negative tuples are collected.

Emotions. Shaver et al. (1987) identify 135 base words which belong to six primary emotion classes: `Anger`, `Fear`, `Joy`, `Love`, `Sadness`, and `Surprise`. Synonyms of the base words are searched manually as extension words [2]. Extension words are classified as the same emotion class as their corresponding base words. We only keep adjectives and verbs. The base and extension words which belong to `Joy` and `Love` are classified as `#emo_pos`, and words from `Anger`, `Fear` and `Sad` are `#emo_neg`. GLM-4 is used to filter out bad cases, and its prompt is listed in Figure 3. Totally 1,425 positives tuples and 1,946 negatives tuples are collected.

3.2 Perceiving States

ASER (Zhang et al., 2022, 2020) is used to extract events from the text of a food review. By resorting to POS tagging [3] and dependency parsing, we detect the following keyword combinations in an event: (1) food entity + feeling state, (2) food entity + emotion state, (3) "I/We" + emotion state, and (4) emotional action. These combinations indicate mental states about food. Keywords link and activate corresponding nodes in Nature Design, like "meatball" and "perfect" in Figure 1 (b). If "not" appears in an event, then feeling and emotion keywords would link and activate opposite nodes. For instance, in the event "I am not happy", "happy" is linked to `#emo_neg` rather than `#emo_pos`.

3.3 Forward Transmitting

Links in a MEA-DAG indicate the direction of signal transmission. The node pointed by a link is tail node, and the node on the other side is head node. Activated nodes in 3.2 send out signals along links to tail nodes, which is a forward transmitting process. For example, in Figure 1 (a), when the node `#emo_pos` is activated, as head node, it sends out a signal which activates its tail node `#need_food_pos`, and then `#action_pos` is activated by its head node `#need_food_pos`.

3.4 Taking Actions

Action events are detected according to patterns listed in Table 2. Only events that have first-person subject "I/We" are considered. Next a `Past` event is determined by checking if the POS tagging of its verb is `VBD` or `VBN`, which is then linked to and activates the node `#past_experience`. Other events, `Modal/Present/Future`, are further classified into three types: `Mental`, `Physical` and `Social` by GLM-4. Definitions and prompts are listed in Figure 3. Events of each type are linked to the following activated nodes respectively, `#mental_action`, `#physical_action` or `#social_action`.

4 Evaluation

4.1 Error Analysis

Totally 92,990 valid MEA-DAGs are extracted out from 568,454 reviews. A valid MEA-DAG is defined as only `#need_food_pos` or

[2]Synonym Website: https://www.merriam-webster.com/thesaurus

[3]Penn Treebank POS tags. Check more details in https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

Table 2: Action event patterns. Only first-person subject patterns are considered. We refer to the ASER pattern writing format.

Pattern	Example
I/We -nsubj-v_1	"I freeze"
I/We -nsubj-v_1-dobj-n_2	"I slice the loaf"
I/We -nsubj-v_1-xcomp-a	"I feel hungry"
I/We -nsubj-(v_1-iobj-n_2)-dobj-n_3	"I give this product 5 star"
I/We -nsubj-v_1-xcomp-a_1-cop-be	"I expect to be served"
I/We -nsubj-v_1-xcomp-n_2-cop-be	"I want to be a gourmet"
I/We -nsubj-v_1-xcomp-v_2-dobj-n_2	"I wait to pay the product"
I/We -nsubj-v_1-xcomp-v_2	"I want to cook"
I/We -nsubj-v_1-nmod-n_2-case-p_1	"I go into the kitchen"
(I/We -nsubj-v_1-dobj-n_2)-nmod-n_3-case-p_1	"I push the pizza into the oven"

#need_food_neg is activated. From valid MEA-DAGs, 100 samples are randomly chosen as a test set for manual evaluation of correctness. A MEA-DAG is incorrect if the revealed relationship is not logically making sense. During the evaluation, seven types of errors are found, whose distribution is shown in Table 3. Totally 42 errors are detected and 37 samples having at least one error. Event Linking Loss and ASER Extraction Loss are the top two error types on the test set. In this section, we discuss why each error type occurs and possible solutions to improve it. Detailed MEA-DAG examples of each error type are appended in Appendix A.

Event Linking Loss. A MEA-DAG fails to incorporate critical information of events, making the revealed relationship not logically complete. The reasons lie in the coarseness of Nature Design, capturing very limited concepts. The limited number of patterns for linking events and nodes also leads to this loss. Besides adding more nodes and patterns, one interesting direction of improvement is to equip the algorithm with learning ability, being able to automatically build new nodes and links when it perceives new events and their outcomes.

ASER Extraction Loss. ASER fails in extracting out critical events from review texts, breaking the logic completeness. This happens due to the limited patterns of event-extractions by ASER. For instance, "I do not know how I could say whether or not the cat food is tasty" would be extracted as one event, "the cat food is tasty", missing key phrases "do not know" and "whether or not". It's necessary to find a method of understanding a sentence as a whole.

Wrong Subsequent Action. An action event should not be linked to the children of #action_pos or #action_neg, as it's not driven by human's need. For example, in the event "I consider myself a pro when it comes to popcorn", the reason of being a pro is not triggered by one specific satisfaction of food-related need, but the rich experience of eating popcorn. This kind of error happens due to lack of deep semantic understanding of an event. Adding more temporal nodes to Nature Design could help to improve the accuracy.

Word Sense Ambiguity. A word is linked to a wrong node due to sense ambiguity. For example, in "you are going to get a light coffee", "light" is incorrectly linked to #emo_pos, as "light" here describes the flavor of coffee. Our methods have no ability to determine which sense of a word should be used in an event. A possible cure might be that MEA-DAGs are built for each sense of a word, and then MEA-DAG merging is implemented between sense and context. A proper sense could be merged smoothly into the context.

Wrong Belief. A word is wrongly linked to emotional or feeling nodes in Nurture Belief. For instance, the words "different" and "raw" are incorrectly connected to #experience_feeling_pos. This happens as SentiWordNet has classification errors, which could not be thoroughly filtered out by LLM. This type of error brings up an interesting research topic: automatic error-correction mechanism. In the course of human development, numerous beliefs are established about this world, some of which are false. Through subsequent experiences, they consistently reinforce correct beliefs and fix wrong beliefs. This mechanism should work perfectly for this kind of error.

Wrong Past Action. Although an action event happens in the past, they are actually driven by how well the food need is met. For instance, "I had to take one star off" is triggered by the dissatisfaction of need. Judging a past action only by tense is not enough. Adding more temporal nodes to Nature Design could help this situation.

Negation Loss. Events are linked to wrong nodes due to failure of capturing negation. "It just failed to deliver" expresses a negative meaning of action, which is hard to capture unless the semantic meaning of "fail" is incorporated in Nature Design. One solution is adding a layer of nodes which is

		Test		Short-Test		Long-Test	
		Count	Percent	Count	Percent	Count	Percent
Sample	Incorrect	37	37.0%	14	29.8%	23	43.4%
	Total	100	100.0%	47	100.0%	53	100.0%
Error	Event Linking Loss	9	21.4%	3	20.0%	6	22.2%
	ASER Extraction Loss	7	16.7%	2	13.3%	5	18.5%
	Wrong Subsequent Action	6	14.3%	1	6.7%	5	18.5%
	Word Sense Ambiguity	6	14.3%	3	20.0%	3	11.1%
	Wrong Belief	6	14.3%	4	26.7%	2	7.4%
	Wrong Past Action	5	11.9%	1	6.7%	4	14.8%
	Negation Loss	3	7.1%	1	6.7%	2	7.4%
	Total	42	100.0%	15	100.0%	27	100.0%

Table 3: Accuracy and error types with count and percentage distribution. We present the results for the test set, the short-test set (sentence number < 5) and the long-test set (sentence number $\geq$ 5).

specifically responsible for dealing with negation and other logic operations.

4.2 Review Length Effect

Depending on whether the number of sentences contained in a review is less than 5, the test set is splitted into a short-test set and a long-test set. By comparing the error differences between the short-test and the long-test, we investigate the effect of review length on accuracy.

Table 3 shows the comparison result. Incorrectness rate of the short-test set is 29.8%, while the long-test set has 43.4%, which indicates that our methods are not well-suited for processing lengthy reviews. Top three error types of the short-test are Wrong Belief, Word Sense Ambiguity and Event Linking Loss, while the long-test are Event Linking Loss, ASER Extraction Loss and Wrong Subsequent Action. From the shift of top three errors, we find that the main bottlenecks of processing long reviews lie in lack of rich nodes and links in Nature Design, as well as lack of comprehensive and in-depth understanding of a sentence.

5 Conclusion

We compute MEA-DAGs to understand the relationships among motivations, emotions and actions from natural language texts. Nature Design is novelly introduced to imitate human's nature, and Nurture Belief connects outside world and human's nature. Our methods are white-box and don't rely on huge annotation resources. Error analysis is implemented to identify the main problems and find possible directions for further optimization.

References

Stefano Baccianella, Andrea Esuli, Fabrizio Sebastiani, et al. 2010. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Lrec*, volume 10, pages 2200–2204.

Edward L Deci and Richard M Ryan. 2000. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, 11(4):227–268.

Deepanway Ghosal, Siqi Shen, Navonil Majumder, Rada Mihalcea, and Soujanya Poria. 2022. Cicero: A dataset for contextualized commonsense inference in dialogues. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5010–5028.

Team GLM, Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Diego Rojas, Guanyu Feng, Hanlin Zhao, Hanyu Lai, Hao Yu, Hongning Wang, Jiadai Sun, Jiajie Zhang, Jiale Cheng, Jiayi Gui, Jie Tang, Jing Zhang, Juanzi Li, Lei Zhao, Lindong Wu, Lucen Zhong, Mingdao Liu, Minlie Huang, Peng Zhang, Qinkai Zheng, Rui Lu, Shuaiqi Duan, Shudan Zhang, Shulin Cao, Shuxun Yang, Weng Lam Tam, Wenyi Zhao, Xiao Liu, Xiao Xia, Xiaohan Zhang, Xiaotao Gu, Xin Lv, Xinghan Liu, Xinyi Liu, Xinyue Yang, Xixuan Song, Xunkai Zhang, Yifan An, Yifan Xu, Yilin Niu, Yuantao Yang, Yueyan Li, Yushi Bai, Yuxiao Dong, Zehan Qi, Zhaoyu Wang, Zhen Yang, Zhengxiao Du, Zhenyu Hou, and Zihan Wang. 2024. Chatglm: A family of large language models from glm-130b to glm-4 all tools. *Preprint*, arXiv:2406.12793.

Lin Gui, Ruifeng Xu, Dongyin Wu, Qin Lu, and Yu Zhou. 2018. Event-driven emotion cause extraction with corpus construction. In *Social Media Content Analysis: Natural Language Processing and Beyond*, pages 145–160. World Scientific.

Carroll E Izard. 1992. Basic emotions, relations among emotions, and emotion-cognition relations.

Abraham Harold Maslow. 1943. A theory of human motivation. *Psychological review*, 50(4):370.

Abraham Harold Maslow. 1958. A dynamic theory of human motivation.

Julian John McAuley and Jure Leskovec. 2013. From amateurs to connoisseurs: modeling the evolution of user expertise through online reviews. In *Proceedings of the 22nd international conference on World Wide Web*, pages 897–908.

George A Miller. 1995. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

Nasrin Mostafazadeh, Aditya Kalyanpur, Lori Moon, David Buchanan, Lauren Berkowitz, Or Biran, and Jennifer Chu-Carroll. 2020. Glucose: Generalized and contextualized story explanations. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4569–4586.

Jaak Panksepp. 2004. *Affective neuroscience: The foundations of human and animal emotions*. Oxford university press.

Soujanya Poria, Navonil Majumder, Devamanyu Hazarika, Deepanway Ghosal, Rishabh Bhardwaj, Samson Yu Bai Jian, Pengfei Hong, Romila Ghosh, Abhinaba Roy, Niyati Chhaya, et al. 2021. Recognizing emotion cause in conversations. *Cognitive Computation*, 13:1317–1332.

Hannah Rashkin, Maarten Sap, Emily Allaway, Noah A Smith, and Yejin Choi. 2018. Event2mind: Commonsense inference on events, intents, and reactions. *arXiv preprint arXiv:1805.06939*.

Maarten Sap, Ronan Le Bras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A Smith, and Yejin Choi. 2019. Atomic: An atlas of machine commonsense for if-then reasoning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 3027–3035.

Phillip Shaver, Judith Schwartz, Donald Kirson, and Cary O'connor. 1987. Emotion knowledge: further exploration of a prototype approach. *Journal of personality and social psychology*, 52(6):1061.

Hongming Zhang, Xin Liu, Haojie Pan, Haowen Ke, Jiefu Ou, Tianqing Fang, and Yangqiu Song. 2022. ASER: towards large-scale commonsense knowledge acquisition via higher-order selectional preference over eventualities. *Artificial Intelligence*, 309:103740.

Hongming Zhang, Xin Liu, Haojie Pan, Yangqiu Song, and Cane Wing-Ki Leung. 2020. ASER: A large-scale eventuality knowledge graph. In *WWW*, pages 201–211.

A Error Examples

A.1 Event Linking Loss

Figure 4 shows two examples whose MEA-DAGs lose critical information of events, resulting that the revealed relationship is not complete in logic sense.

A.2 ASER Extraction Loss

Figure 5 presents two examples in which ASER couldn't extract critical events from review texts. As a result, the generated MEA-DAG is incomplete.

A.3 Wrong Subsequent Action

Figure 6 presents two examples in which action events are wrongly linked to the children nodes of `#action_pos`. If the contexts and timeline of events are considered, they should be linked to `#past_experience`.

A.4 Word Sense Ambiguity

Figure 7 presents two examples in which words are wrongly linked to feeling or emotion nodes, as our methods have no ability to determine the sense of a word given its context.

A.5 Wrong Belief

Figure 8 presents two examples in which events are linked to incorrect nodes due to errors in Nurture Belief.

A.6 Wrong Past Action

Figure 9 presents two examples in which events are wrongly linked to `#past_experience`, as they are not factors that affect whether the need is met. In fact, they are the result of food need not being met.

A.7 Negation Loss

Figure 10 presents two examples in which events are wrongly linked to `#emo_pos`, as the negation expressions "don't", "whether or not" and "doesn't" are not captured by our methods.

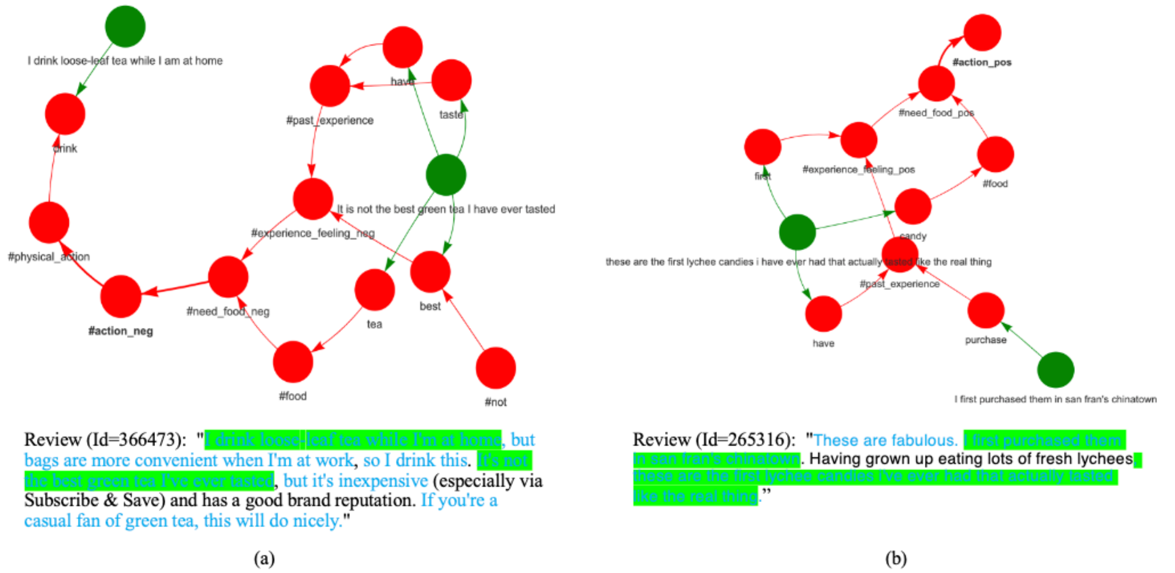


Figure 4: Events extracted by ASER are in blue color. We use a green font background to highlight the events incorporated in MEA-DAG. (a): Critical events "bags are more convenient when I'm at work", "it's inexpensive" are not included in MEA-DAG. (b): Critical event "these are fabulous" are not included in MEA-DAG.

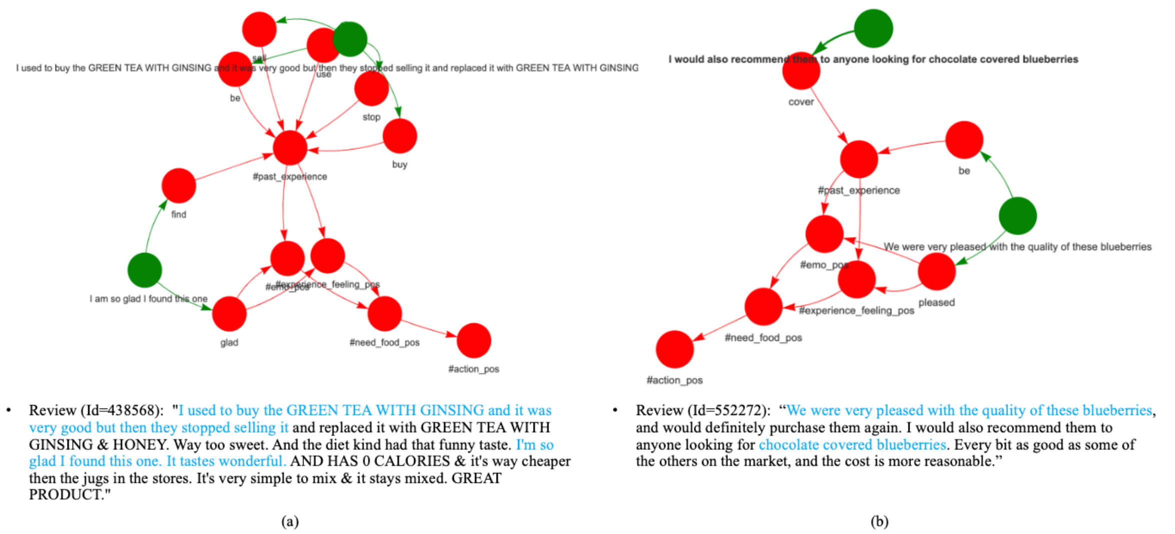
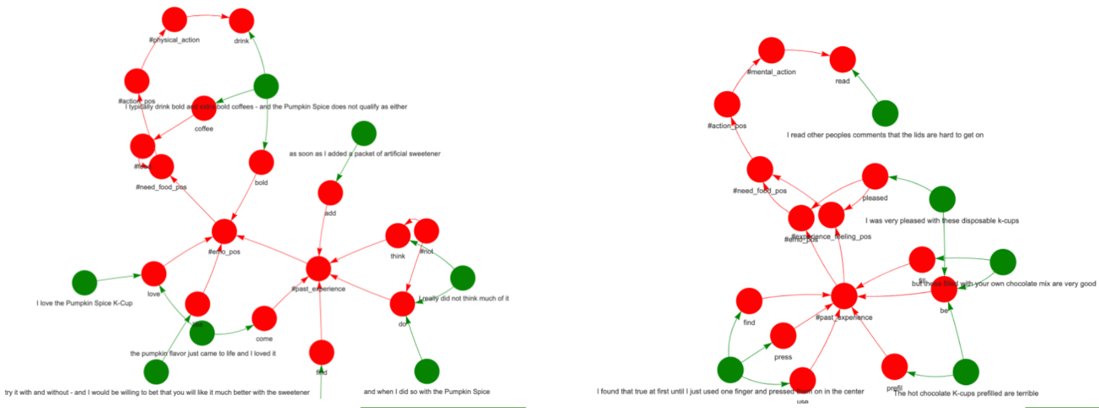


Figure 5: Texts in blue color are the events extracted by ASER. (a): Critical events "the diet kind had that funny taste", "it's way cheaper then the jugs in the stores" and "it's very simple to mix & it stays mixed" are not captured by ASER. (b): Critical events "would definitely purchase them again", "I would also recommend them" and "the cost is more reasonable" are not captured by ASER.



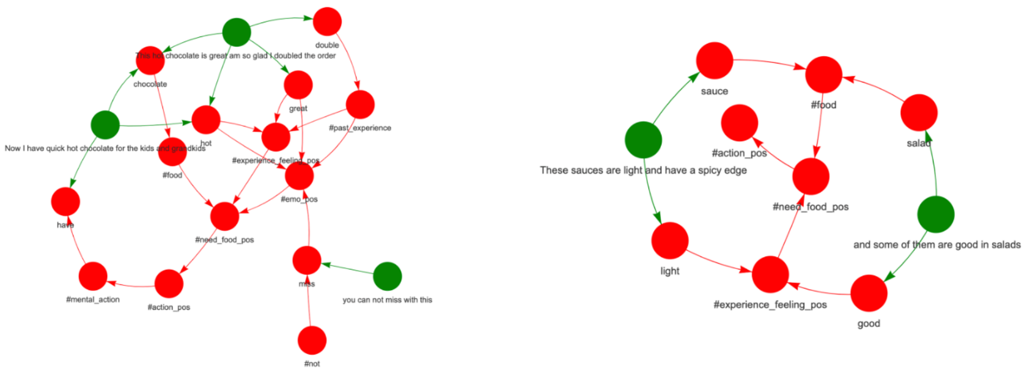
- Review (Id=112029): "I love the Pumpkin Spice K-Cup. I typically drink bold and extra bold coffees - and the Pumpkin Spice doesn't qualify as either. But it smells just like pumpkin pie. and tastes pretty close to it as well. One thing I found out though is to definitely add some kind of sweetener (whether it be artifical sweetener or sugar). I always drink coffee with just cream. and when I did so with the Pumpkin Spice, I really didn't think much of it. …"

(a)

- Review (Id=6993): "I was very pleased with these disposable k-cups. I read other peoples comments that the lids are hard to get on. I found that true at first until I just used one finger and pressed them on in the center. I heard a click and they were ready to go. I will order a larger amount next time to save on shipping. The hot chocolate K-cups prefilled are terrible, but these filled with your own chocolate mix are very good."

(b)

Figure 6: Examples with Wrong Subsequent Action error. We use a green font background to highlight the wrong events. (a): "I typically drink bold extra bold coffee" is linked to `#physical_action`. However, it's not driven by `#need_food_pos`, as the word "typical" indicates that it's a habitual action. (b): "I read other peoples comments that the lids are hard to get on" is linked to `#mental_action`. However, it should be linked to `#past_experience`. Considering its context, "read" in this event is in the past tense, representing an action that occurred in the past.



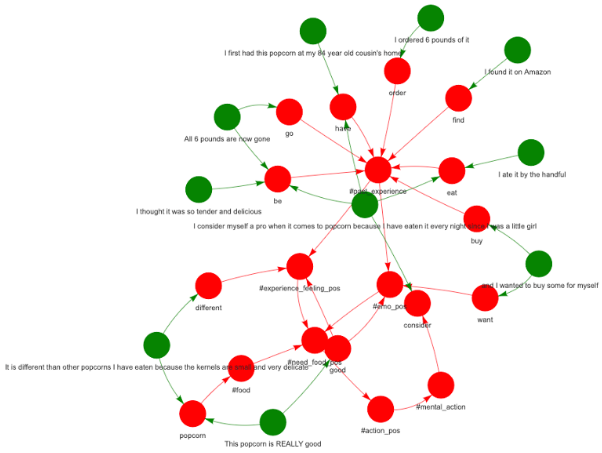
- Review (Id=567786): "This hot chocolate is great am so glad I doubled the order, you can not miss with this. Now I have quick hot chocolate for the kids and grandkids."

(a)

- Review (Id=73971): "These sauces are light and have a spicy edge. They mix well with meats, Mexican food, veggies, and some of them are good in salads. They are liquid, so blend well with other foods. Jalapeno is the lightest, Spicy Ranchero is like the original Taco Bell Hot Sauce Flavor the most. It goes well as a salad dressing if applied sparingly. The Chipotle one is the most mellow one, all with a bite. I will definitely buy again. Taco Bell has outdone themselves. "

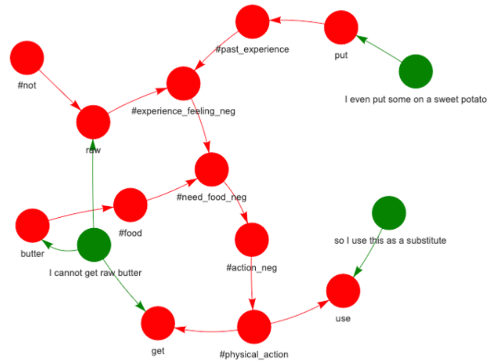
(b)

Figure 7: Examples with Word Sense Ambiguity error. We use a green font background to highlight the wrong events. (a): In the event "I have quick hot chocolate for the kids and grandkids", "hot" is an objective description of food, not bearing an positive emotion or feeling sense. (b): In the event "these sauces are light and have a spicy edge", "light" is wrongly linked to `#experience_feeling_pos`, as it describes sauce taste, not a feeling.



- Review (Id=2717): "I first had this popcorn at my 84 year old cousin's home. I thought it was so tender and delicious, and I wanted to buy some for myself. I found it on Amazon. com believe it or not. Anyway, I ordered 6 pounds of it. All 6 pounds are now gone. This popcorn is REALLY good. It is different than other popcorns I have eaten because the kernels are small and very delicate. I consider myself a pro when it comes to popcorn because I have eaten it every night since I was a little girl. I usually eat my popcorn 1 kernel at a time, but since the kernels are so small, I ate it by the handful. My only complaint is that it does cost a lot to have it shipped. I wish I could find it locally so I would not have to pay for the shipping costs. If you like popcorn, I highly recommend it."

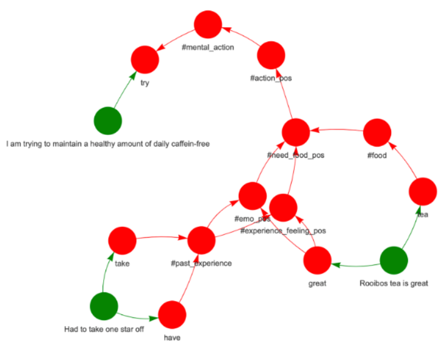
(a)



- Review (Id=11657): "Great to make popcorn with. I even put some on a sweet potato, I cannot get raw butter, so I use this as a substitute."

(b)

Figure 8: Examples with Wrong Belief error. We use a green font background to highlight the wrong events. (a): The word "different" is incorrectly linked to `#experience_feeling_pos`. (b): The word "raw" is incorrectly linked to `#experience_feeling_pos`.



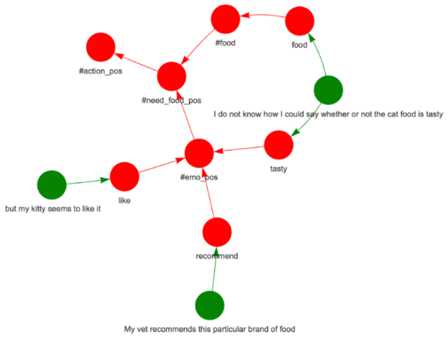
- Review (Id=265043): "I'm trying to maintain a healthy amount of daily caffein-free, sugar-free liquid intake, but plain water gets a bit boring after a while. Rooibos tea is great: fragrant, convenient, and it's even supposed to be anti-oxidant. The only problem is that the packaging of this product is ridiculously annoying. The plastic is very difficult to peel. I had to struggle with it for like 5 minutes, even with the help with my swiss army knife. Had to take one star off."

(a)

- Review (Id=179467): "This tea is really tasty. I wasn't even a tea person until tasting this, and once I did, I was hooked. The flavor is light and super-refreshing but you can still taste and smell the mint. The only reason why I took off one star is because I tend to like things that taste strongly, and this one is a little bit delicate for me. To remedy that, I either steep it with hibiscus or black tea, or I steep it at double strength and all is well. Putting it with honey and stevia also tastes great, and makes me feel better when I have colds. I love this tea."

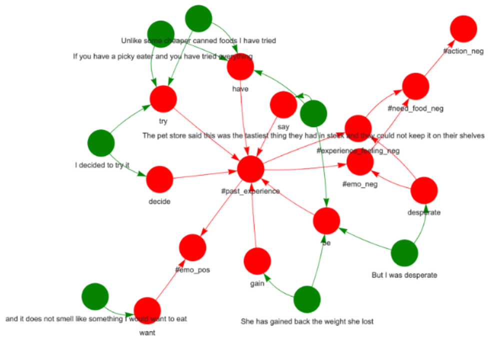
(b)

Figure 9: Examples with Wrong Past Action error. We use a green font background to highlight the wrong events. (a): The event "Had to take one star off" happened due to dissatisfaction with food. Therefore, it should be linked to `#physical_action`. (b): The event "The only reason why I took off one star " is driven by dissatisfaction with food, not a factor that affects whether the need is met.



- Review (Id=369849): "I don't know how I could say whether or not the cat food is tasty, but my kitty seems to like it. My vet recommends this particular brand of food, so that's good enough for me."

(a)



- Review (Id=489837): "I was hesitant to buy this because it costs a fortune and venison and fish sounded pretty unappetizing. But I was desperate, because I have a 16 year old dog who's always been a picky eater but had now lost a dangerous amount of weight because of a heart condition. The pet store said this was the tastiest thing they had in stock and they couldn't keep it on their shelves, so despite its $4/can price tag, I decided to try it. The food looks gross. It has strange-looking pieces in it (tripe, some exotic part of the venison or blue mussel.), and it doesn't smell like something I would want to eat, either. But my dog cannot get enough of it. She gobbles it up and begs for more. She's gained back the weight she lost, and feeding her is no longer an excruciating, hour long process of coaxing and pleading. Unlike some cheaper canned foods I've tried, it does not give her digestive problems. If you have a picky eater and you've tried everything, try Ziwipeak. Maybe your dog just has a sophisticated palette :-)."

(b)

Figure 10: Examples with Negation Loss error. We use a green font background to highlight the wrong events. (a): The event "I don't know how I could say whether or not the cat food is tasty" is wrongly linked to #emo_pos due to failure of capturing "not". Although "tasty" has a sense of positive emotion, "not" changes its meaning to opposite side. (b): The event "it doesn't smell like something I would want to eat" is wrongly linked to #emo_pos, as "doesn't" is not captured in the MEA-DAG.