

LREC 20
Palma **26**



Disambiguation of Emotion Annotations by Contextualizing Events in Plausible Narratives



LREC 2026

Johannes Schäfer and Roman Klinger

Fundamentals of Natural Language Processing,
University of Bamberg, Germany



Motivation: Emotion Analysis

Example event/sentence

A person describes an **experience** as follows:



“The loudspeaker suddenly malfunctioned and went silent.”

wanted to listen

person responsible

Emotion analysis task: Guess the **emotion** this event evoked in the person.

Examples of emotion categories: anger, boredom, disgust, fear, guilt, joy, pride, relief, sadness, shame, surprise, trust, no-emotion.

→ **Several emotions are possible in certain contexts!**

Example in Different Contexts

1. I was tasked with giving a presentation to a large crowd.
2. **The sound system malfunctioned, amplifying my voice to an ear-piercing level.**
3. The sound technician ignored the problem and chatted with someone.
4. The **audience covered their ears** and looked at me with discomfort.

Event: *The loudspeaker suddenly malfunctioned and went silent.* → **Relief**

1. I arrived at a remote wilderness survival training camp, where the instructors emphasized the **importance of following loudspeaker instructions** for safety.
2. The instructors warned us about the toxic waste site nearby and explained that the loudspeaker would alert us to any changes in air quality.
3. During the first exercise, I struggled to navigate the challenging terrain, but the **loudspeaker provided crucial guidance**, helping me stay on track.
4. I completed a difficult obstacle course, relying heavily on the loudspeaker's instructions to avoid hazards and find the safest route.

Event: *The loudspeaker suddenly malfunctioned and went silent.* → **Fear**

What this example shows



Event descriptions often have ambiguous emotion interpretation

1. Different labels can result from different **perspectives**.

→ Several works explain label variations by studying differences between human annotators, e.g. different demographics/traits/values.

2. Little research on variations due to **missing information**. Humans typically can explain **their interpretation** by providing background information.

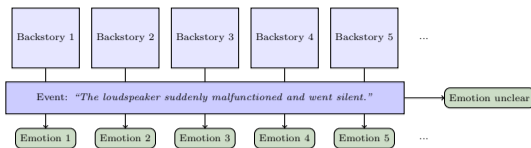
→ [This paper](#):

- Can we use an LLM to **generate context** resulting in a specific interpretation?
- How are different **emotions** influenced by **contexts**?

Overview



Research goals



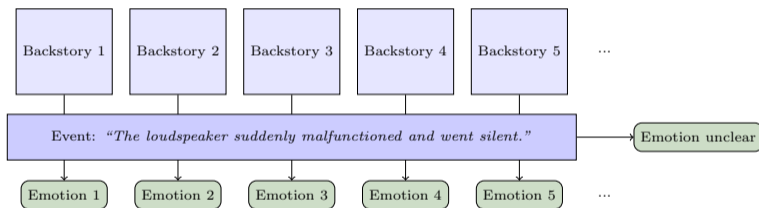
- **Specialized dataset:** Set of event descriptions
→ each with different backstories resulting in different emotions (13 categories)
- **Analysis:** Do the different emotions become clearer when including backstory?

Necessary steps

1. Develop backwards generation method (given last event of narrative)
2. Data quality analysis (automatic and human annotation)
3. **Measure changes in emotion analysis when including backstory**

Influence of Generated Context in Emotion Analysis

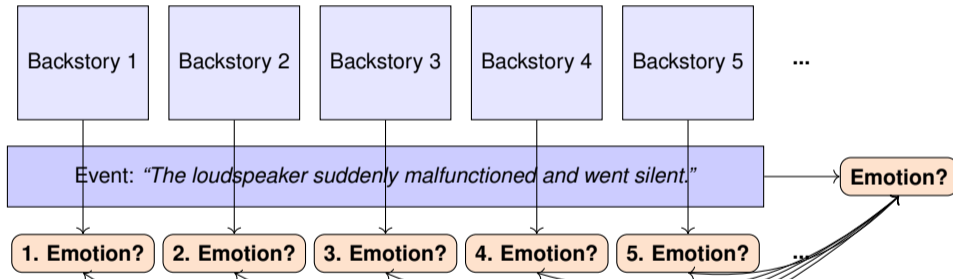
Dataset: 1,000 event descriptions, each with 13 backstories (4 sentences)
→ 13,000 narratives



- Inter-Annotator Agreement in human annotations
- LLM-predicted emotion probabilities

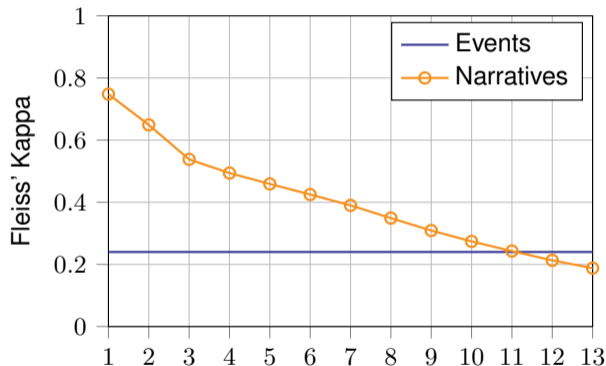
Human Annotation Setup

Annotators assign an emotion category given different input:



13 comparisons: Annotation agreement on event vs. on narrative?

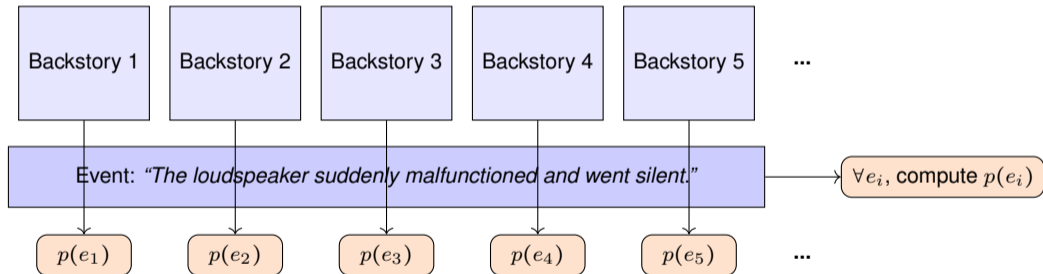
Human Inter-Annotator Agreement (IAA)



- IAA improves on average in 10 out of 13 comparisons, when adding a backstory to an event
- ⇒ Generated contexts make **emotions clearer** for human annotators

Setup for Automatic Emotion Analysis

Use LLM to compute emotion label probabilities



→ Evaluation: Averages over 1,000 instances, or 1,000 narratives per emotion

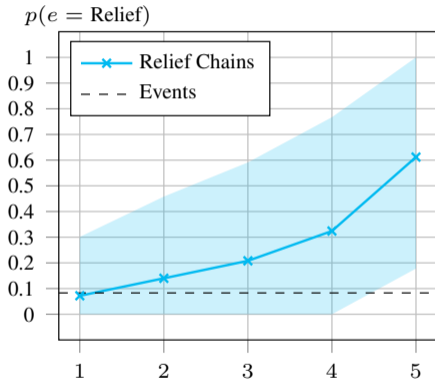
Automatic Emotion Analysis Results



	Anger	Boredom	Disgust	Fear	Guilt	Joy	Pride	Relief	Sadness	Shame	Surprise	Trust	No-Emot.
Events	.02	.01	.00	.01	.00	.33	.01	.08	.01	.00	.14	.02	.37
Narratives	.24	.12	.24	.28	.24	.27	.24	.61	.42	.27	.30	.13	.09

- High prior probability for joy/no-emotion → not increased when narrative is included
 - Other emotion categories: **Predicted probabilities increase substantially**
 - Strongest **influence of context**: relief, sadness, fear
- ⇒ **System recognizes contextual disambiguation**

Emotion Trajectories in Relief Narratives



Relief example:

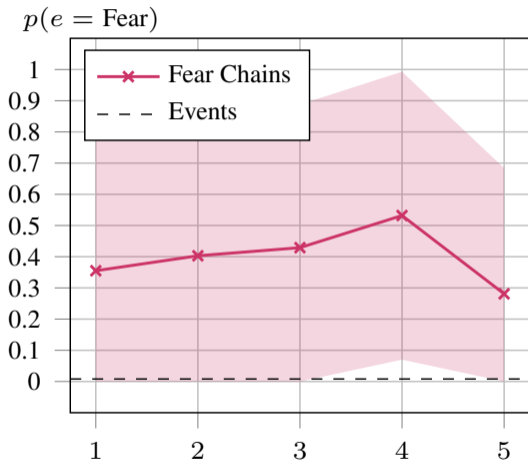
1. I was tasked with giving a presentation...
2. The sound system malfunctioned, amplifying my voice to an ear-piercing level.
3. The sound technician ignored the problem ...
4. The audience covered their ears ...

Event: *The loudspeaker suddenly malfunctioned and went silent.*

Here: Average over 1000 narratives for Relief

- Steady increase throughout narratives
 - Strongest surge on introduction of last event
- ⇒ Relief strongly relies on what happens in last event, can clearly be influenced by narrative

Emotion Trajectories in Fear Narratives



- Steady increase from 1-4; decline in probability upon introduction of last event
 - Extremely low prior probability
→ Last events generally not aligned with this emotion
- ⇒ Fear is mainly built by the backstory

Conclusion



Context plays a crucial role in emotion analysis

- **Comprehensive dataset** allows for **systematic** analysis and training of models: 13,000 narratives; for each of 1,000 event descriptions, one backstory per emotion
- Generated **backstories** make **potential interpretations** of readers specific
- Method can be used to **provide context which explains certain interpretations** (e.g. in XAI), show alternative interpretations
- Different **emotions** are evoked differently in narratives of event sequences