

MINERAL : Multi dimensional representation for conversational modeling

Bechet Frédéric¹, Damnati Géraldine², Favre Benoit¹, Fourtassi Abdellah¹, Guinaudeau Camille³, Ghannay Sahar³, Grouin Cyril³, Lecorvé Gwénolé², Rosset Sophie³, Shimorina Anastasia²

¹Laboratoire Informatique et Systèmes, ²Orange Innovation, ³Laboratoire Interdisciplinaire des Sciences du Numérique

Challenges in Spoken Conversation Analysis

- Non-verbal information can be essential to understand the message,
- Visual grounding between participants is often needed during a conversation to integrate posture and body gesture as well as references to the surrounding world,
- Semantic grounding between participants of a conversation to establish mutual knowledge is essential for communicating with each other

==> Use conversation transcription as a textual input of a generative model for downstream tasks can provide helpful results

==> But still limited: summary only convey the spoken content, regardless of the other dimensions

Objective

Train a multimodal conversation representation model for communicative acts (i.e., the smallest—verbal or non-verbal—communication units with a consistent communication intention)

Study communicative structures (how and why communicative acts are linked),

Generate an enriched transcriptions in the form of a script

Exploit a panel of ecological conversational datasets from various domains to evaluate the quality of the generated scripts and their impact on relevant use cases



[Amy and Sheldon are speaking to Amy's mother over a video computer connection on a laptop]

Mrs. Fowler : It's nice to meet you too, Sheldon. I honestly didn't believe Amy when she told me she had a boyfriend. (to Sheldon Cooper)

Sheldon Cooper : I assure you I am quite real. And I'm having regular intercourse with your daughter. (to Mrs. Fowler)

Mrs. Fowler : [in a surprised tone] What?

Sheldon Cooper : Oh, yes. We're like wild animals in heat. It's a wonder neither of us has been hurt. (to Mrs. Fowler)

Mrs. Fowler : [scared] Amy, what is he saying? (to Amy Farrah Fowler)

Amy Farrah Fowler : You wanted me to have a boyfriend, mother. Well, here he is. (to Mrs. Fowler)

[Sheldon waves at the computer screen, while Mrs. Fowler nervously waves back]

Research questions

- ◆ How to group low-level descriptors into communication acts with the same underlying intention?
- ◆ What types of relationships between communication acts need to be captured to reflect the structure of a conversation?
- ◆ Which representation of conversation structure is most conducive to tasks involving automatic comprehension of a conversation?
- ◆ What are the dimensions of interest for representing communication acts for script generation?
- ◆ Does the structure in communication acts contribute to the explainability of the generation process?
- ◆ Can fiction training corpora contribute to producing reproducible results in real-world conditions?
- ◆ To what extent are scripts a useful representation format for solving different practical cases?

Organization

WP1: preparation of multimodal data

WP2: representation and extraction of communicative acts

WP3: representation and extraction of relations between CAs

WP4: script generation

WP5: analysis of how scripts and representations from the previous WPs can help for specific use cases

