Multimodal Machine Learning Lab

Winter Semester 2025/2026

Niklas Deckers and Martin Potthast

Deep Semantic Learning
University of Kassel and hessian.AI

Agenda

- Encoder-Decoder Scenarios

- Spanning a Space of Supported Tasks

- Identifying System Components

- Adversarial Scenarios

Encoder-Decoder Scenarios

❑ Each game iteration will consist of a single encoder-decoder scenario, involving a task, a sender (working on the encoding step), one or multiple receivers (working on the decoding step), and a verification step (which may involve human feedback)

❑ The components of the system should be replaceable

❑ An experiment has access to the data produced by such iterations and controls the general setup (which task is used, number of receivers, etc.)

Spanning a Space of Supported Tasks

- ❑ Objective: Find abstract requirements to a general system
- ❑ Define supported tasks and their characterizing properties
- ❑ Quick, Draw!: Drawing and multiple choice
- ❑ Emoji Movie Trivia: Restricted form set
- ❑ Dixit: Complex verification
- ❑ 4 Pics 1 Word: Multiple images
- ❑ ESP Game: Encoding by writing a text
- ❑ Skribbl.io: Multiple decoders

Identifying System Components

- ❑ Encoding method:

    - Provides the user with images and/or text as specified by the task

    - Allows the user to input: free text (optionally restricted by constraints or a time limit); drawing; a selection of elements from a restricted form set (emojis, vocabulary, drawing arrows for *transactional structures* as described by Gunther Kress and Theo van Leeuwen in *Reading Images*); or some multiple choice or ordering

- ❑ Decoding method:

    - Provides the user with the given encoded message

    - Allows the user to input one of the modalities specified above

- ❑ Verification method:

    - Could use measures such as text comparison (BLEU, Levenshtein distance, etc.); multiple choice comparison (correctness verification, nDCG for rankings); checking pre-specified criteria (like in Portrayal)

    - May also consult humans

    - Should also work with multiple receivers

Identifying System Components

- ❏ Tasks:

    - Specifies instructions shown for each of the components
    - Restrictions for the choice of encoding/decoding/verification methods
    - Lists of examples with values given to the encoding/decoding/verification methods
    - Some AI to drop in for the sender or the receiver

Competitive Scenarios

- ❏ Two receivers working against each other (on the same task)
- ❏ A sender and a receiver working against a third party (a receiver): Either by giving the sender hints about the legitimate receiver's identity and background (similar to designing questions for intergenerational trivia such as Mind the Gap [solidroots.com]) or by having a human work against the machine (where the objective is to encode hints in such a way that only a human would understand, forming a sort of Reverse Turing Test)
- ❏ For the latter, the sender could also be a machine, which would resemble a machine-generated Captcha
- ❏ Some games could also call for multiple senders (e.g., Dixit), but this will not be explored further here since it makes it unclear which (and whose) information provided the most valuable hint