Introduction to Dialogue Systems

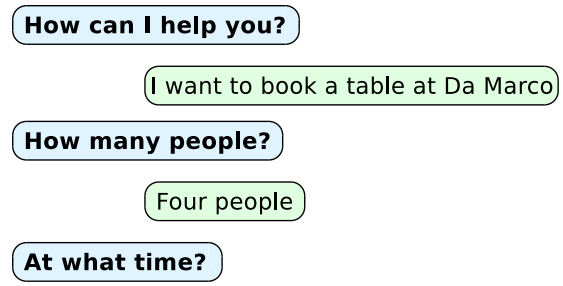
UNIVERSITY OF
GOTHENBURG

CHALMERS

Richard Johansson

richajo@chalmers.se

dialogue systems

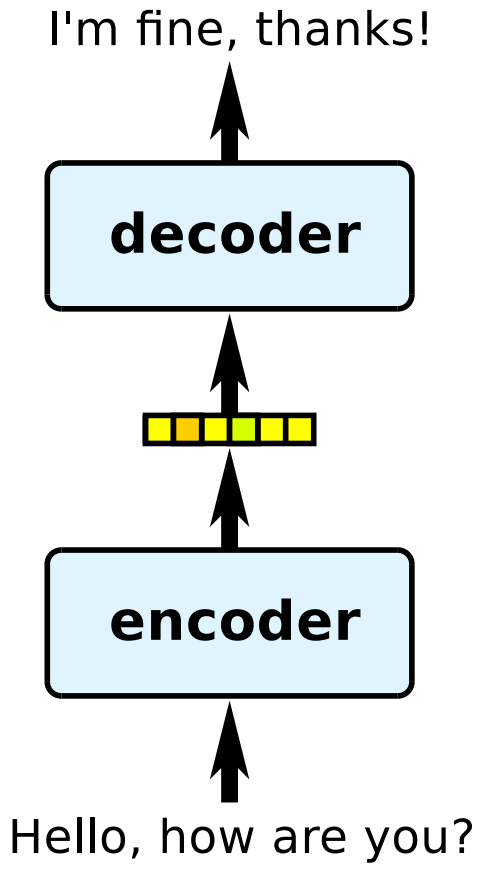


- "chatbots", "conversational agents", ...
- computational systems that we can **interact** with
- general challenges:
  - representing **state**
  - interacting with **external resources**
- no discussion of speech technologies here! ASR/TTS at endpoints

types of dialogue systems

- chit-chat / open-domain:
  - mostly for entertainment value
  - some use in therapy, education
  - often unrestricted in terms of content
- task-oriented:
  - searching, booking, troubleshooting, ...
  - often more restricted/"controlled"

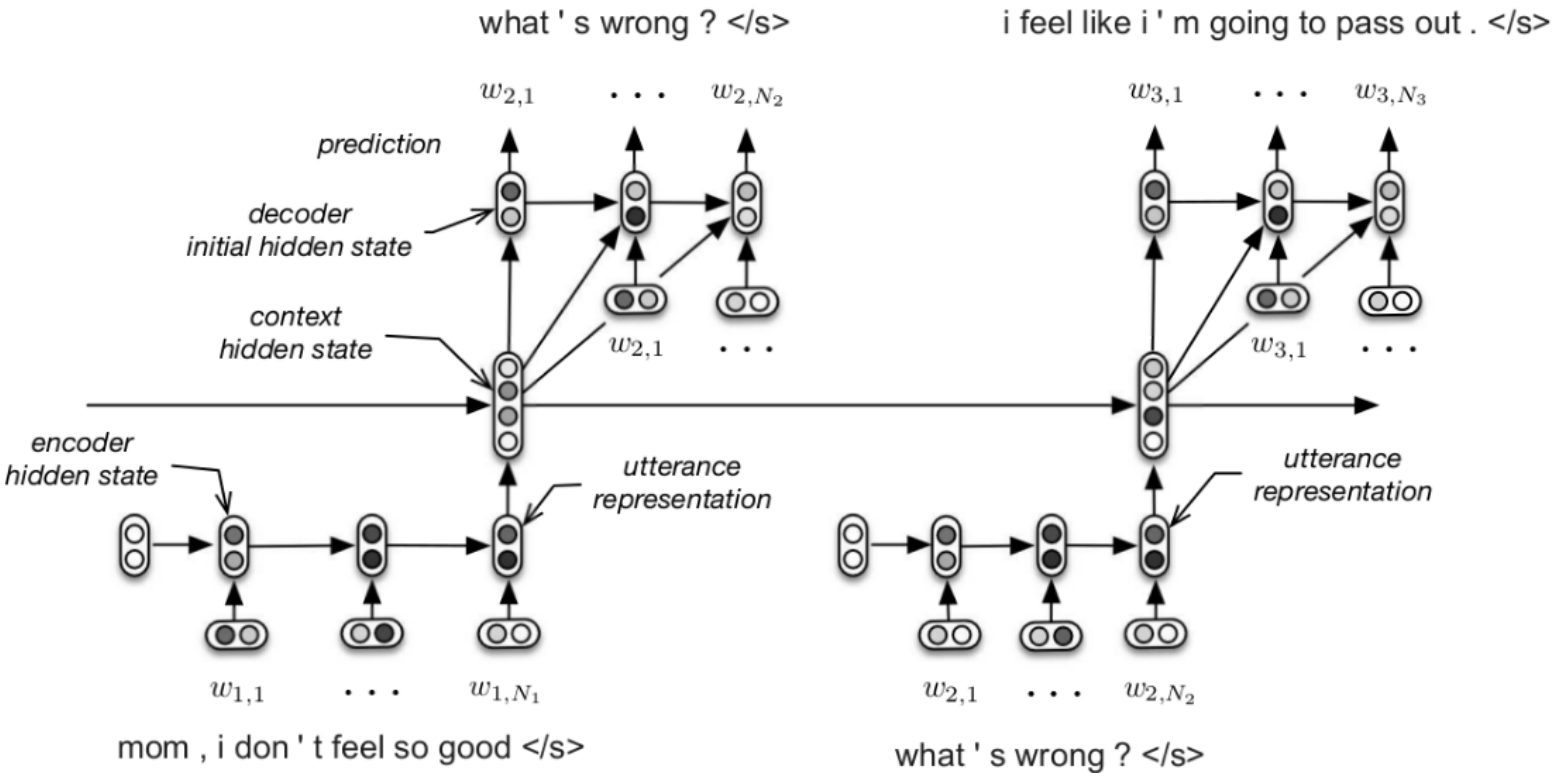
open-domain response generation: basic idea



- basic idea introduced by Vinyals and Le (2015)

how do we represent the **state** (the **history**) of the conversation?

architecture with turn representations



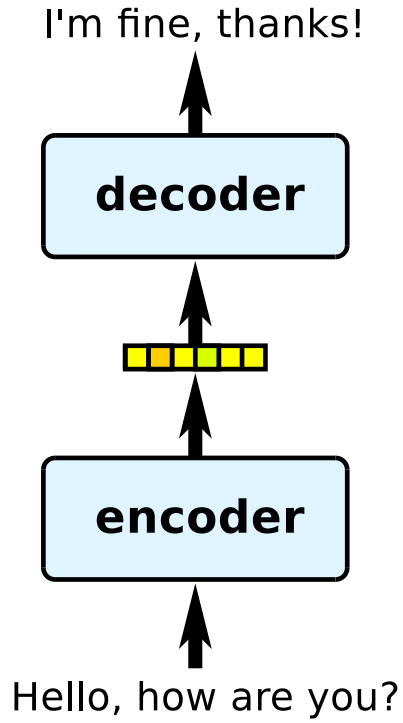
(Serban et al., 2016)

example: DIALOGPT

- the **DIALOGPT** model (Zhang et al., 2020) is a GPT-2 transformer tuned on Reddit conversations
- like the original GPT models, the response generation task is framed as **language modeling**: a Transformer decoder
- the dialogue history is used as a **prompt** (start of sequence)

examples: Meena and Blender Bot

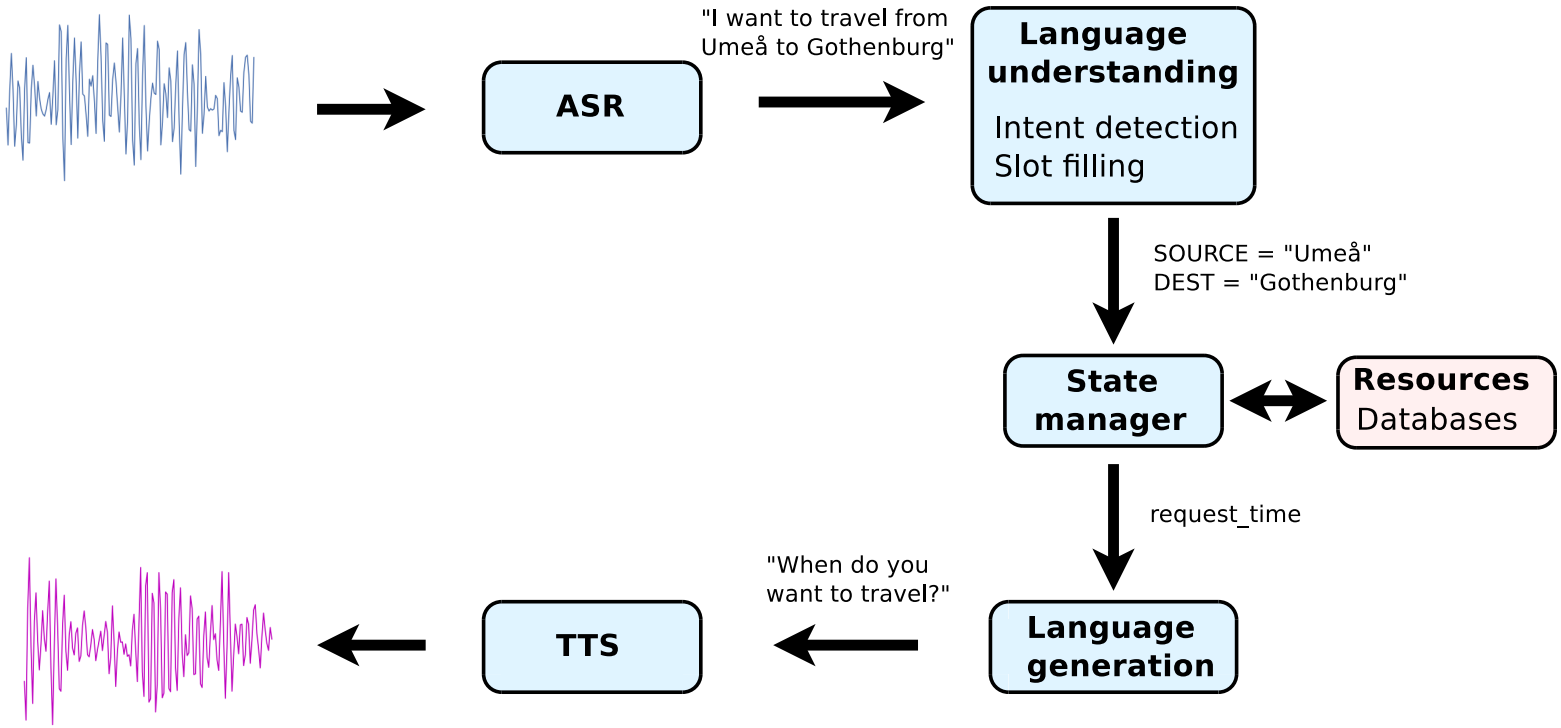
- Google's **Meena** (Adiwardana et al., 2020) and Facebook's **Blender Bot** (Roller et al., 2021) use encoder/decoder Transformers: the dialogue history is the input



open research problems in open-domain dialogue

- **decoding** from the language model: avoiding bland answers
- maintaining **consistency over a long conversation**: "goldfish memory" (Xu et al., 2021)
- **consistency with the outside world** (Komeili et al., 2021)
- **avoiding toxicity** in generated outputs
- Huang et al. (2020) gives an overview of challenges

traditional task-oriented dialogue systems

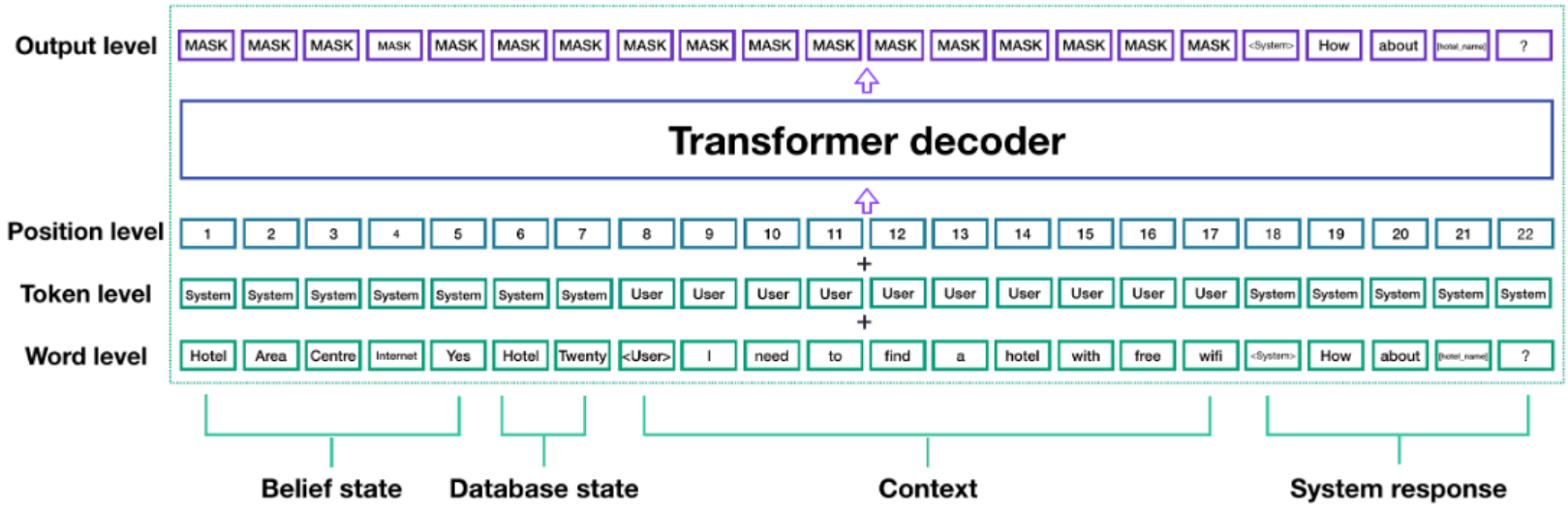


task-oriented dialogue: challenges

- domain-specific design requires **modularity**
  - research has often focused on a module and less on end-to-end performance
  - pipeline design $\Rightarrow$ risk of error propagation
- need to interact with **external resources**
- often highly specialized, training data often scarce

adapting pre-trained models for task-oriented dialogue

- Budzianowski and Vulic (2019) adapt GPT-2 for task-oriented response generation



- **TOD-BERT** (Wu et al., 2020) is a BERT-like model that could be used for several steps in task-oriented dialogue

dialogue datasets

- Reddit, Twitter, ...
- subtitles (http://opus.nlpl.eu/OpenSubtitles-v2018.php)
- technical support e.g. Ubuntu (Vinyals and Le, 2015)
- ConvAI2 (Dinan et al., 2020) http://convai.io/
- Persona-Chat (Zhang et al., 2018) https://github.com/facebookresearch/ParlAI/tree/main/parlai/tasks/personachat
- overview: *A Survey of Available Corpora for Building Data-Driven Dialogue Systems* (Serban et al., 2018)

evaluating dialogue systems

- word-based: overlap, perplexity, ...
- human-oriented evaluations: user satisfaction, fluency, ...
- there are conflicting opinions on how well they correlate:
  - Dinan et al. (2020) report weak correlations in the ConvAI2 challenge
  - Adiwardana et al. (2020) report good correlations between perplexity and human scores for the Meena chatbot

references

D. Adiwardana, M.-T. Luong, D. So, J. Hall, N. Fiedel, R. Thoppilan, Z. Yang, A. Kulshreshtha, G. Nemade, Y. Lu, and Q. V. Le. 2020. Towards a human-like open-domain chatbot. arXiv:2001.09977.

P. Budzianowski and I. Vulic. 2019. Hello, it's GPT-2 - how can I help you? towards the use of pretrained language models for task-oriented dialogue systems. In *Proceedings of the 3rd Workshop on Neural Generation and Translation*.

E. Dinan, V. Logacheva, and V. Malykh et al. 2020. The second conversational intelligence challenge (ConvAI2). In *The NeurIPS '18 Competition*.

M. Huang, X. Zhu, and J. Gao. 2020. Challenges in building intelligent open-domain dialog systems. *ACM Trans. Inf. Syst.* 38(3).

M. Komeili, K. Shuster, and J. Weston. 2021. Internet-augmented dialogue generation. arXiv:2107.07566.

S. Roller, E. Dinan, N. Goyal, D. Ju, M. Williamson, Y. Liu, J. Xu, M. Ott, E. M. Smith, Y.-L. Boureau, and J. Weston. 2021. Recipes for building an open-domain chatbot. In *EACL*.

I. Serban, R. Lowe, and P. Henderson et al. 2018. A survey of available corpora for building data-driven dialogue systems. *Discourse & Dialogue* 9(1):1–49.

I. V. Serban, A. Sordoni, and Y. Bengio et al. 2016. Building end-to-end dialogue systems using generative hierarchical neural network models. In *AAAI*.

O. Vinyals and Q. Le. 2015. A neural conversational model. arXiv:1506.05869.

C.-S. Wu, S. C. Hoi, R. Socher, and C. Xiong. 2020. TOD-BERT: Pre-trained natural language understanding for task-oriented dialogue. In *EMNLP*.

J. Xu, A. Szlam, and J. Weston. 2021. Beyond goldfish memory: Long-term open-domain conversation. arXiv:2107.07567.

S. Zhang, E. Dinan, J. Urbanek, A. Szlam, D. Kiela, and J. Weston. 2018. Personalizing dialogue agents: I have a dog, do you have pets too? arXiv:1801.07243.

Y. Zhang, S. Sun, M. Galley, Y.-C. Chen, C. Brockett, X. Gao, J. Gao, J. Liu, and B. Dolan. 2020. DIALOGPT: Large-scale generative pre-training for conversational response generation. In *ACL*.