Artificial Emotional Intelligence: Dialogue Systems in Medicine

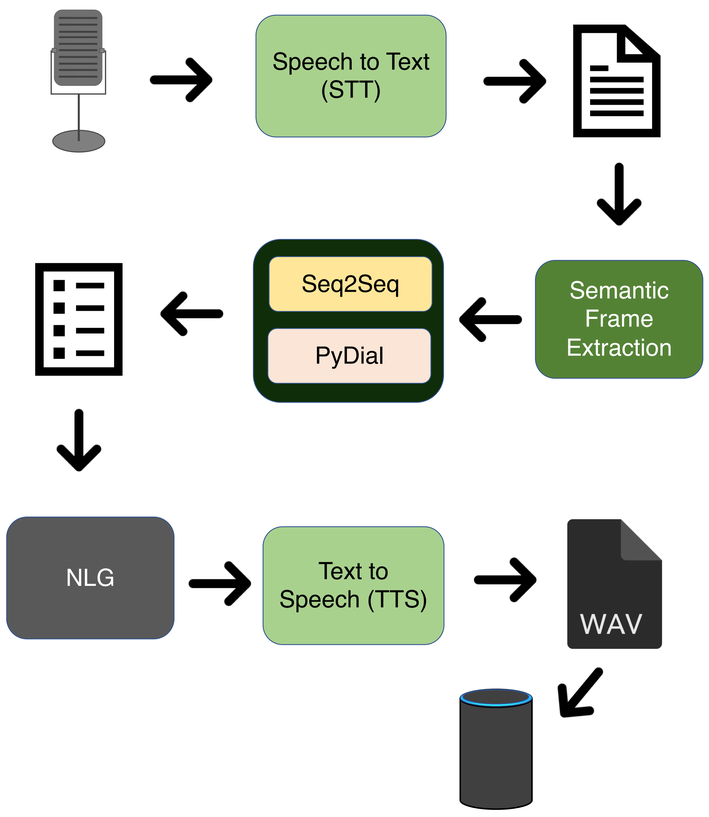
Vishnu Arun, Thomas Huffstutler, Raymond Finzel, Libby Ferland, Serguei Pakhomov (PhD), Maria Gini (PhD)
Department of Computer Science and Engineering, University of Minnesota

College of Science & Engineering
University of Minnesota
Driven to Discover℠

INTRODUCTION

- This research was conducted in response to the growing gap between the design considerations of commercial dialogue state tracking systems and older users (55+) or users with mild cognitive impairment.
- Through this research, we developed a human-computer interface for our dialogue agent using a Raspberry Pi. Our current dialogue agent uses a rule-based model, however, we are developing a Sequence-To-Sequence model to increase generalization ability to non-specific dialogue tasks.
- Using natural language processing we seek to develop a dialogue agent that can comprehend medical terminology and distinguish it from casual conversation.
- **Natural Language Processing** is a technique where a machine attempts to process and interpret human language. An end-to-end dialogue system requires two components: An NLG unit and an NLU unit.
- **NLG** (Natural Language Generation) is a module that can generate natural language from a semantic frame.
- **NLU** (Natural Language Understanding) is a module that can extract semantic understanding from natural language.
- **Semantic Frames** are a collection of the characteristic attributes or coherent structures of a dialogue that can be encoded into an organized human-readable file such as a JSON.

ARCHITECTURE



ISSUES

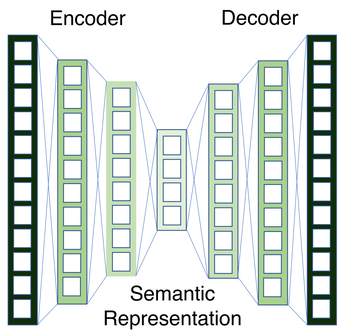
How will data privacy be protected?

- On-device vs. Cloud
- Personalized recommendations vs. a general purpose model

Ethical Issues regarding AI in the Healthcare domain.

- Trusting an AI to provide confidential information
- AI incorporating empathy into its responses
- Ensure consistently high accuracy

MODELS

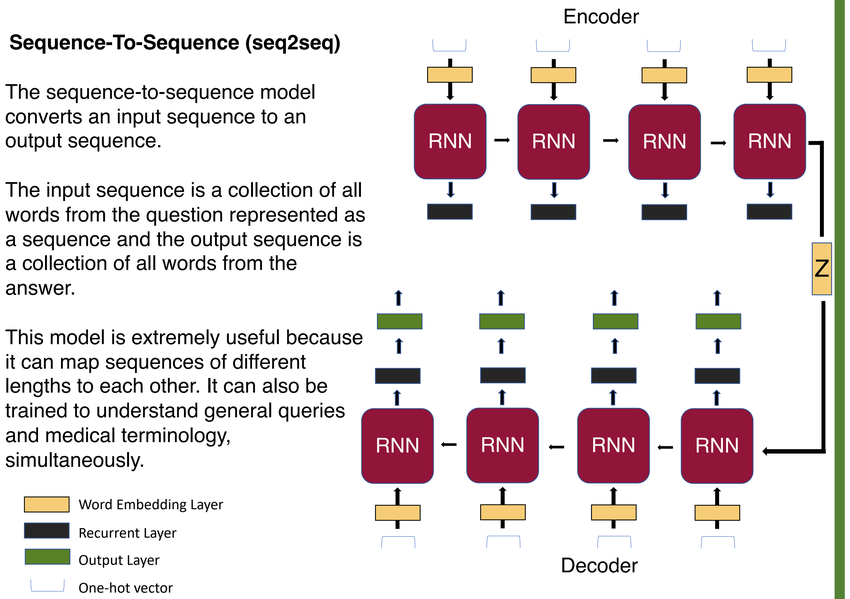


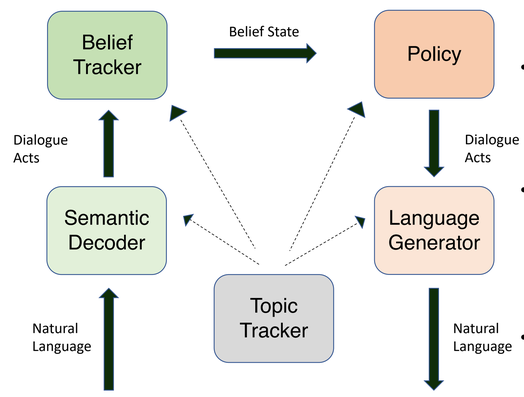
Autoencoder

- An autoencoder is a type of ANN that aims to copy its input to its output.
- It compresses the input into a latent-space representation then attempts to reconstruct the original input.
- We attempt to learn a more meaningful representation of the input data that extracts the semantical structure.

Sequence-To-Sequence (seq2seq)

- The sequence-to-sequence model converts an input sequence to an output sequence.
- The input sequence is a collection of all words from the question represented as a sequence and the output sequence is a collection of all words from the answer.
- This model is extremely useful because it can map sequences of different lengths to each other. It can also be trained to understand general queries and medical terminology, simultaneously.

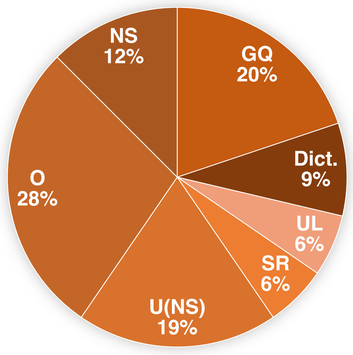




Rule Based Models

- A rule-based model employs hard-coded rules in order to understand a query and generate a response.
- These models can capture specific phenomena very easily, but become cumbersome when attempting to generalize the system.
- The PyDial Framework is an example of a rule-based model.

EXPERIMENTAL RESULTS



NS: Not Sure
GQ: General/Complex Questions
O: Other
U(NS): Understanding (non-specific)
SR: Speech Recognition
UL: Unusual Language or Names
Dict.: Text Dictation

Figure 1. This chart shows the most frequent difficulties with smart devices as collected from a D2D survey conducted on respondents 55+.

SR: Speech Recognition
LU: Language Understanding
UNS: Understanding (non-specific)
SP: Security and Privacy
LC: Learning Curve/Knowledge Gap
O: Other
NS: Not Sure

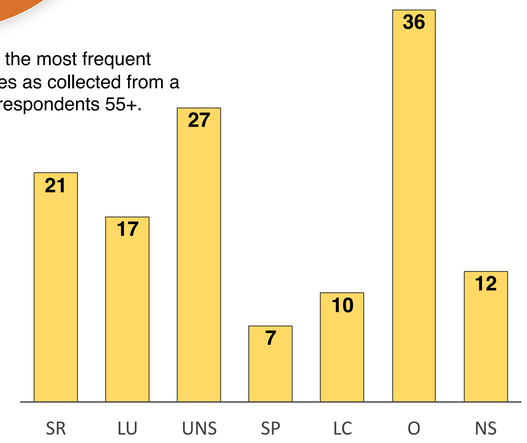


Figure 2. Results of D2D survey show the most undesirable qualities of current commercially available dialogue state tracking systems.

- This data was collected at the 2018 Minnesota State Fair by the Next Generation Robotics group through a research survey.
- The proposed use cases for patients with mild cognitive impairment should take into account the lack of specialized medical knowledge for audiences and ease of usability for unfamiliar users.
- These results show that in the medical domain it is important to keep the baseline of knowledge required of the user to a minimum.
- Dialogue systems in the medical domain should be able to distinguish medical terminology from non-medical terminology from context of interaction.

FUTURE WORK

- We are looking to develop a human computer interaction component to the system so that patients can more naturally interact with it.
- We are also working on incorporating an autoencoder to improve semantic frame construction and representation compression.

REFERENCES

1. Li, Xiujun & Panda, Sarah & Liu, Jingjing & Gao, Jianfeng. (2018). Microsoft Dialogue Challenge: Building End-to-End Task-Completion Dialogue Systems.
2. Zheng, Joan & Finzel, Raymond & Pakhomov, Serguei & Gini, Maria. (2018). Spoken Dialogue Systems for Medication Management, *W3PHIAI 2019*, AAAI, 2019.
3. Ferland, Libby & Huffstutler, Thomas & Rice, Jacob & Zheng, Joan & Ni, Shi & Gini, Maria. (2018). "Evaluating Older Users' Experiences with Commercial Dialogue Systems: Implications for Future Design and Development ," *DEEP DIAL 2019*, AAAI, 2019.
4. Li, Jiwei, et al. "A Hierarchical Neural Autoencoder for Paragraphs and Documents." *ArXiv:1506.01057 [Cs]*, June 2015. *arXiv.org*, http://arxiv.org/abs/1506.01057.