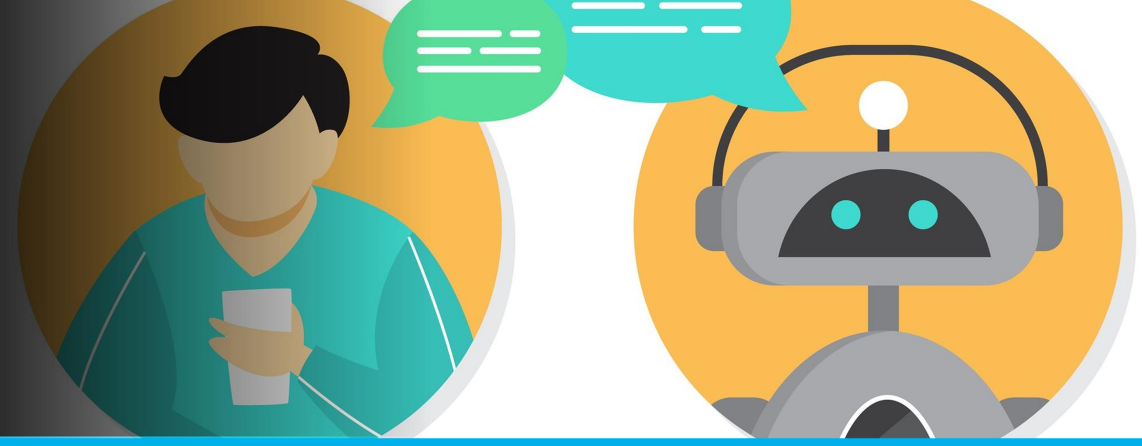

The Open Domain Interviewing Agent

Ming-Bin Chen | Professor. Michael Witbrock | School of Computer Science at the University of Auckland

Problem description

Interviewing is a critical human-to-human task in multiple industries; It is an iterative exploring process that adjusts accordingly based on the responses and the context. This project aims to devise a technique to **create an interviewing agent that can ask multi-turns questions exploratively and strategically to maximize information gain.**

Background and related study

- Most of the question generation techniques developed have **answers available in context.**
- Mao Nakanishi et al. (2019) proposed a question generation methodology which identifies focal context segments and question patterns to generate questions. Meanwhile, Scialom et al. (2019) also have achieved answer-agnostic question generation through data augmentation and reinforcement learning. **Both approaches only generate single turn questions.**
- Most recently, Qi et al. (2020) **set the baseline for open-domain conversational information-seeking question generation** by building a reinforcement learning process involving a teacher agent with target information, and a student agent for question generation. **The question generation are rewarded for better specificity and informativeness of the questions.**

Interviewing Strategy

- While some conversational attributes, including fluency, relevance etc. are essential, some should be adaptive accordingly to the context and history. The reinforcement learning process should aim to optimize full session information gain than single turn attributes, such as specificity.
- This project hopes to train the interviewing agent to pickup strategies such as:
  1. **Avoiding/changing topics** when negative signals have been received, e.g. the interviewee already has responded "Don't know".
  2. Asking questions that have **the best potential of retrieving important information**.
  3. **Adapting the specificity level depending on the conversation history**. Ask a deeper question when a certain trend is detected or ask a more general question when there is lack of information.
  4. **Avoiding asking the same or similar questions** that expect the same or similar answers.

THE UNIVERSITY OF AUCKLAND
Te Whare Wananga o Tamaki Makaurau
NEW ZEALAND

SCIENCE
SCHOOL OF COMPUTER SCIENCE
MACHINE LEARNING

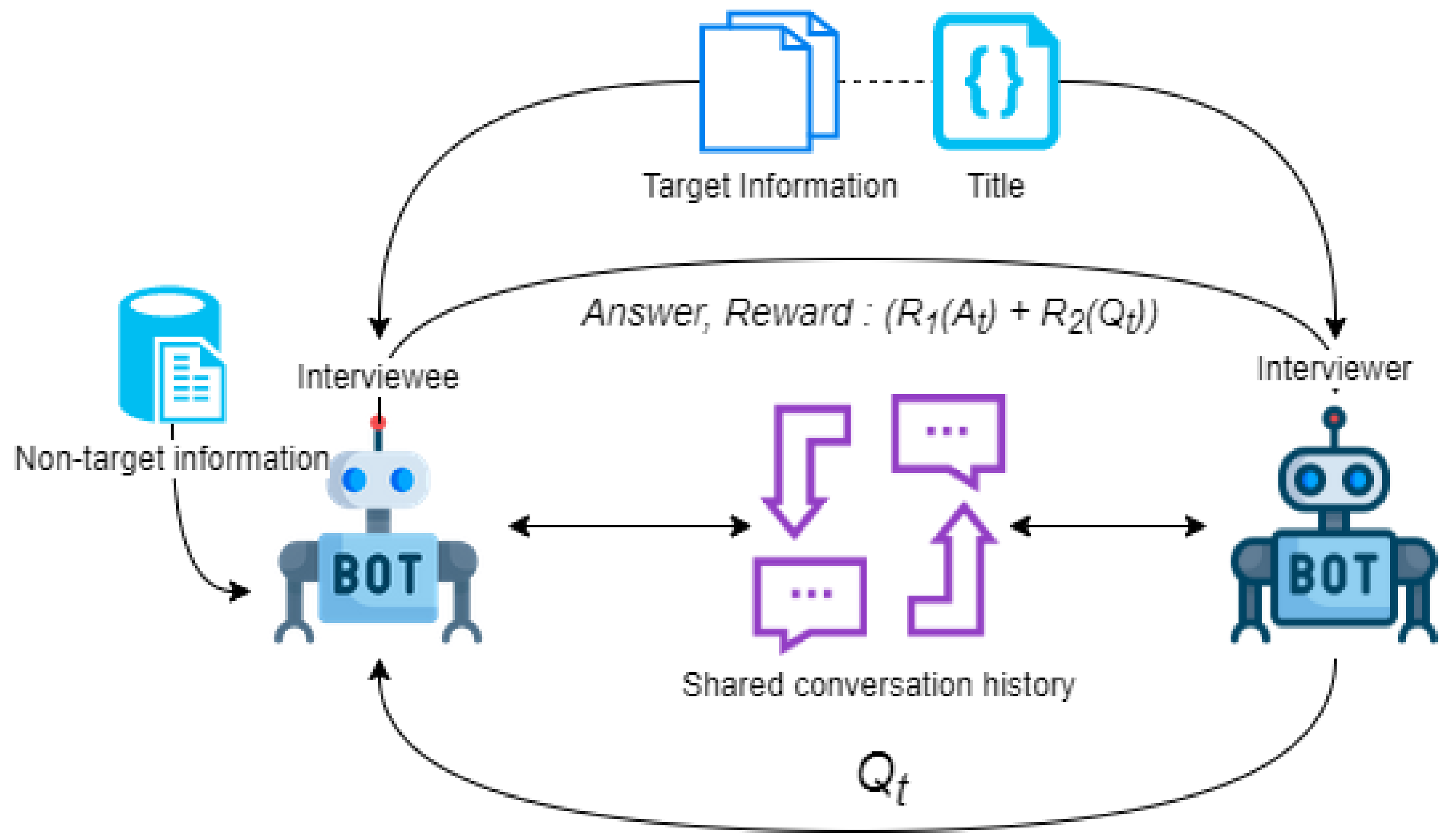

Figure 1: A brief view of the reinforcement learning between the interviewee(QA model) and interviewer (QG model) with shared history. A target text is given to the interviewee, while the title is given to the interviewer. Occasional non-target information is introduced as noise. The reward is based on both the questions and the answers.

Three Stages Training Method

Interviewee model preparation	Answer agnostic question generator pretraining	Reinforcement learning for interviewing strategy
• The Quac (Choi et al, 2018) dataset is used to train a multi-turns, context based QA model as interviewee.	• The Quac dataset is also used to pretrain the QG model by only using the titles, answers and questions parts of the dataset. • Test the effect of data augmentation on other dataset.	• Freeze the parameter of the interviewee • Learn to adjust **weight decoding** to change specificity and relatedness. • Optimize interviewer for local features with R_1 and global information gain with R_2

Weighted decoding

- Using technique suggested by See et al (2019), the interviewer is allowed to effectively control the **specificity, relatedness, and repetition** of the generated questions by adding additional weights during the decoding process to affect the output vocabulary selection. This project uses reinforcement learning to find the best weights in a given context and history. ***Wi*** is the weight for i-th attribute, given the attribute scoring function ***fi***, vocabulary **w**, input ***x.*** and previous sequence ***y***.

$$\sum_i w_i * f_i(w; y_{<t}, x).$$

Reinforcement rewards

- The reinforcement reward for a given question answering turn ***t*** is the sum of local reward R_1 and global reward R_2, given generated question Q_t and interviewee response A_t, and the history conversation $C_{<t}$:
  $$R_t(Q_t, A_t; C_{<t}) = R_1(Q_t; C_{<t}) + R_2(A_t; C_{<t})$$
- R_1 is the sum of a group of linguistic quality scoring functions, such as fluency, relevance etc.
- R_2 is the global reward for the overall information gain. Qi et al (2020) has a similar reward function that measures the number of the non-overlapping unigrams in the current answer to the most similar answer in the history as informativeness. Here the project takes some more sophisticated steps:
  - A vector of importance scores S = $\{I_1, I_2, \dots\dots, I_n\}$ for every token in the target text T = $\{W_1, W_2, \dots\dots, W_n\}$ are precalculated by either normalised tfidf/ Bert multi-head attentions/ Longformer global attention. The sum of the vector is 1.0.
  - At the step ***t***, The importance function $Imp(C_t)$ is therefore the sum of all weighted unique index tokens captured in the history. And the R_2 is the difference of the history before and after the current update.
  - $R_2 = Imp(C_{<t} + A_t) - Imp(C_{<t})$

Proposed evaluation methods

- Traditional automatic matrices, including perplexity, and BLEU.
- Measuring the importance score weighted sum of the final history.
- We also evaluate the correctness of the retrieved information by running the QA model with the context text replaced by the retrieved QA history.
- The generated conversation is evaluated by human following the formats of Qi et al (2019). In addition, the strategic interviewing pattern are also observed and recorded.

To be continued

This is still an ongoing project, currently yet to have significant results to report. This project aims to contribute by:

- Improving Qi et al (2020) work by adapting the global reward function, to make the interviewer more importance-orientated.
- Validating the uses of weighted decoding in the scenario of explorative question generation.
- Experimenting which importance scoring function works better.
- Devising and validating new evaluation methods for this emerging field.

Reference

- Nakanishi, Mao, Tetsunori Kobayashi, and Yoshihiko Hayashi. 2019. Towards Answer-unaware Conversational Question Generation. In Proceedings of the 2nd Workshop on Machine Reading for Question Answering, pp. 63-71
- Scialom, Thomas, and Jacopo Staiano. 2019. Ask to Learn: A Study on Curiosity-driven Question Generation. arXiv preprint arXiv:1911.03350.
- Qi, Peng, Yuhao Zhang, and Christopher D. Manning. 2020. "Stay Hungry, Stay Focused: Generating Informative and Specific Questions in Information-Seeking Conversations." arXiv preprint arXiv:2004.14530.
- See, Abigail, Stephen Roller, Douwe Kiela, and Jason Weston. "What makes a good conversation? How controllable attributes affect human judgments." In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 1702-1723. 2019.
- Choi, Eunsol, He He, Mohit Iyyer, Mark Yatskar, Wen-tau Yih, Yejin Choi, Percy Liang, and Luke Zettlemoyer. 2018. "Quac: Question answering in context." arXiv preprint arXiv:1808.07036.