Ondřej Plátek

Charles University in Prague
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics
Malostranské námestí 25
11800 Prague 1, Czech Republic

oplatek@ufal.mff.cuni.cz
http://ufal.mff.cuni.cz/ondrej-platek

1 Research Interests

My research focuses on automating task-oriented dialogue system deployment. In particular, I believe that end-to-end neural conversation agents will substantially simplify bootstrapping dialogue systems in narrow domain. Having a working dialogue system which can be optimized jointly using single objective function is crucial for my further research where I want focus on improving the dialogue system interactively from conversations.

The task of improving a dialogue systems over time has gained focus of many researchers. However, most of the works focus only on optimizing a dialogue manager with very limited action space (Gašić et al., 2011). Such system is able to optimize only action selection despite its most severe problem may lie in language understanding or natural language generation. Reinforcement learning of end-to-end system updates the system parameters so the most severe errors according to objective function are mitigated first.

The quality of task oriented systems depend not only how well they can communicate with user about their expertise - their domain, but to larger extent it depends also on quality of the database. A common problem even with high quality databases is that they easily obsoletes and the stored facts are no longer valid or incomplete. To mitigate the problem we would like to train a conversational agent to collect missing information from users.

Currently, I focus on simplifying the architecture of dialogue systems using neural networks models so one can train decent baseline from fewer data. So far, I have implemented dialogue-state tracking model (Plátek et.al, 2016) which uses encoder-decoder(Bahdanou et al., 2014) architecture which I would like to extend to end-to-end system. In recent work (**?**), a dataset from restaurant domain is collected with simple annotations which is intended to be used for training end-to-end system. Last but not least, I am interested in rapid prototyping of spoken dialogue systems using current technologies to test simple ideas e.g., how to extract information from users.

1.1 Work in progress: Task oriented End-to-End Dialogue Systems

I am currently working on a prototype of end-to-end dialogue system and its evaluation on Cambridge restaurant domain (Williams et al., 2013). We used the relevant part of the model to estimate the model potential on much easier task of dialogue state tracking (Plátek et.al, 2016) and verified that the model is able to capture dialogue history well. On the other hand, inspired from (Wen et al., 2016) we aim to optimize access to database jointly but without explicitly modeling the dialogue state. We plan to train the model with annotations of API calls instead of dialogue state labels. The work is still in progress, but we plan to evaluate the final model using coverage metrics, which prefer factual answers with similar entities based on the knowledge base ontology.

1.2 Future plans: Extracting annotations from users

In longer perspective, we aim at training the dialogue system so it will be able to collect explicit feedback - annotation interactively from users. We plan to explore how users react to incorrect replies. Later, we intend to detect a situations when user thinks he or she was provided with incorrect reply. Detecting misunderstanding is an important task since it is the first step to obtaining interactively a ground through annotation of the incorrect reply from users.

2 Future of Spoken Dialog Research

I suppose that in near future we will see massive use a very simplistic and narrow domain conversational agents through platforms like Wit.ai or Kick.com. The power of such platforms is the organic selection of domains which are convenient for such simplistic conversations. The obvious challenge is navigating the user to the desired narrow domain and both platforms do not use dialogue to do so effectively ignoring the problem. As as result multiparty dialogue will become important field.

In contrast the virtual assistants such as Siri, Cortana or Google Now will become more open-domain and more conversational agents. I assume that a lot of effort will

be invested in automatic scaling from a few dialogue domains to truly open-domain dialogue.[1] The neural networks model proved to scale well in terms of removing handcrafted components and modeling abstraction layers so the dialogue systems can be easily deployed. On the other hand, knowledge based approaches in question answering perform well and scale from simple domain to open-domain. From my perspective there are at at least two problems which need to be solved before these technologies can be successfully combined. First one need to solve how to parametrize the neural conversational models based on part of the knowledge graph under discussion. Second much more challenging task is to determine what is the knowledge graph under discussion using the neural model and the conversation history.

3 Suggestions for Discussion

My suggestions for the discussion roughly correspond to my research topics:

- End-to-End dialogue systems:
  - How to optimize database access jointly together with conversation?
  - What kind of additional supervision to provide to text-to-text task oriented conversational agents?
  - Is next utterance classification (Lowe et al., 2016) annotation feasible to obtain? Can it be used similarly as noise contrastive estimation (Gutmann, Michael and Hyvärinen, Aapo, 2010) for language modelling?
- Detecting misunderstanding:
  - How to detect misunderstanding automatically?
  - What training data is necessary for misunderstanding detection and how to collect them?
- Rapid prototyping of SDS:
  - Which technology to use?
  - Deployment strategies

References

[Povey et al. 2011] Povey, Daniel and Ghoshal, Arnab and Boulianne, Gilles and Burget, Lukáš and Glembek, Ondřej and Goel, Nagendra and Hannemann, Mirko and Motlíček, Petr and Qian, Yanmin and Schwarz, Petr and others 2011 *The Kaldi speech recognition toolkit*.

[Plátek et.al 2016] Ondˇrej Plátek, Petr Belohlávek, Vojt ˇ ech Hude ˇ cek, and Filip Jur ˇ cí ˇ cek Plátek, Ondřej and Bělohlávek, Petr and Hudeček, Vojtěch and Jurčíček, Filip 2016. *Recurrent Neural Networks for Dialogue State Tracking* To Appear in Slo-NLP 2016 Proceedings.

[Bahdanou et al. 2014] Dzmitry Bahdanau and Kyunghyun Cho and Yoshua Bengio 2014. *Neural Machine Translation by Jointly Learning to Align and Translate* CoRR, abs/1409.0473.

[Gašić et al. 2011] M Gašić, F Jurčíček, B Thomson, K Yu, S Young 2011. *On-line policy optimisation of spoken dialogue systems via live interaction with human subjects* Automatic Speech Recognition and Understanding (ASRU), 2011 IEEE Workshop, 312-317.

[Williams et al. 2013] Williams, Jason and Raux, Antoine and Ramachandran, Deepak and Black, Alan 2013. *The dialog state tracking challenge* Proceedings of the SIGDIAL 2013 Conference, 404-413.

[Wen et al. 2016] Wen, Tsung-Hsien; Vandyke, David; Mrksic, Nikola; Gasic, Milica; Rojas-Barahona, Lina M.; Su, Pei-Hao; Ultes, Stefan; Young, Steve 2016. *A Network-based End-to-End Trainable Task-oriented Dialogue System* Arxiv, http://arxiv.org/abs/1604.04562.

[Lowe et al. 2016] Ryan Lowe and Iulian Vlad Serban and Michael Noseworthy and Laurent Charlin and Joelle Pineau 2016. *On the Evaluation of Dialogue Systems with Next Utterance Classification* Arxiv, https://arxiv.org/abs/1605.05414

[Gutmann, Michael and Hyvärinen, Aapo 2010] Gutmann, Michael and Hyvärinen, Aapo 2010. *Noise-contrastive estimation: A new estimation principle for unnormalized statistical models.* AISTATS

Biographical Sketch

Ondřej Plátek is currently entering third second year as a PhD student in the area of Computational Linguistics at Charles University in Prague. Much of his experience with conversational agents comes from industry research focused not only spoken dialogue systems but also just on automatic speech recognition. He has enjoyed several freelancing opportunities for small companies. His interests include machine learning, linguistics, outdoor sports especially rock-climbing.

[1] I wonder if users truly want to know the information in the large knowledge graphs.