Review of the Ph.D. Thesis

Goal Oriented and Open Domain
Dialogue Management

Candidate: Miroslav Vodolan

The submitted Ph.D. thesis „Goal Oriented and Open Domain Dialogue Management" deals with two relatively independent problems in the construction of dialogue systems, namely goal-oriented and open domain dialogue systems. Let us state that both areas are topical and contain many dissertable problems.

Concerning goal oriented dialogue, the candidate describes a baseline rule-based tracker whose functionality is improved by modifying the speech understanding module and, in particular, by implementing a hybrid approach to monitor the status of the dialogue. To verify the proposed methods Miroslav Vodolan uses data sources that were provided in the State Dialogue Tracking Challenges 2013 and 2014. It is an annotated data from Cambridge restaurant information system allowing the users to change their mind during the conversation. This approach uses machine learning methods to refine the probabilities of transitions between states that were originally obtained by a pure knowledge approach. The resulting solution achieves state-of-the-art parameters. To this part of the thesis I have a question for discussion:

- Candidate describes (in Chapter 5) the dialogue system with the ASR features. Apparently, it is assumed that the ASR system works with a zero OOV (out-of-vocabulary) words. The question is how the system would deal in real dialogue with spoken words outside the dictionary?

The second part of the thesis deals with the design of dialogue system working in open domains. Candidate starts from a proven concept of designing dialogue based on manual work of domain experts and this knowledge-based approach (usually used for designing single domain dialogue) significantly expands and adds new original interactive learning method for the design of open domain dialogue. To verify the proposed approach with in interactive learning, a new dataset of 1900 annotated dialogues between users and system is designed and collected. The candidate uses the created dataset for evaluation of denotation extractions. Denotations are obtained by the application of neural networks, which brings very interesting results. The dialogue management is designed in a new and interesting way and it is shown that by interacting with the user the system learns and can acquire new knowledge for its further operation. This part of the thesis I consider well processed and proving very

good knowledge of the candidate. However, also to this part of thesis I have some questions and comments:

- Although the author offers a theoretical principle of the functioning of a dialogue system on an open domain, it seems to have been implemented and evaluated on only one domain - restaurant search. Does the candidate already have real practical experience with multi-domain dialogues?
- The dialogue system was tested on a limited number (only) text-based concepts. Could the candidate address the question of correctness of the derived facts in the case of the "expected" error rate of ASR and NLU?
- How would a larger (large) number of concepts that need to be evaluated in a multi domain dialogue system affect the system's ability to conduct dialogue in real time?

Conclusion:

In the end of my review I must appreciate a lot of hard theoretic and experimental work, which bring many new knowledge and ideas. The work is also valuable in providing a summary of existing methods and in performing case studies that can guide future work. The candidate proposed the new dialogue state tracker and the original hybrid method of interactive learning for the design of open domain dialogue and verify its correctness on the collected dataset of real dialogues.

Miroslav Vodolan has fulfilled all the objectives planned in the thesis. He has demonstrated the ability to work in research and to achieve scientific results. I recommend the thesis for presentation with the aim of receiving the degree of Ph.D. according to the §47, sect. 4 of the law No. 111/98 (Dig.). The thesis fulfills all conditions of creative work and it is possible to recommend it to the defense.

prof. Josef Psutka, UWB Pilsen

Pilsen, August 15, 2019

As for the goal-oriented dialogue systems dissertation uses data sources that were provided in the State Dialogue Tracking Challenges 2013 and 2014.

Pokud jde o goal riented dialogue systems využívá disertační práce datové zdroje, které byly poskytnuty v Dialogue State Tracking Challengues v letech 2013 a2014. Jedná se o anotovaná data from Cambridge restaurant information system allowing the users to change their mind during the conversation. Kandidát pospisuje (kapitola 3.) baseline rule-based tracker, jehož funkci vylepšuje úpravami modulu porozumění řeči a především nasazením hybridního přístupu ke sledování stavu dialogu. Tento přístup využívá metod strojového učení k zpřesnění prqvděpodobností přechodů mezi stavy, které byly původně získány čistě znalostním přístupem. Výsledné řešení dosahuje state-of-the-art parametrů (výsledků).

K této části práce mám námět do diskuse. V kapitole 5 popisu kadidát dialogový system s ASR features. Zřejmě se předpokládá, že ASR system pracuje s nulovým OOV (out-of vocabulary) slovníkem. Je otázkou, jak by se v reálném provozu dialogový system vypořádal s promluvenými slovy mimo slovník (což je

Candidate vychází z osvědčeného principu designing dialogue based on manual work of domain experts. This knowledge-based approach that is usually use for designing single domain dialogue - disertant navrhuje interactive learning pro návrh open domain dialogue.

Pro ověření navrženého přístupu s interaktivním učením je navržen a collected new dataset of 1900 annotated dialogues between users and system. Vytvořeného datasetu využívá autor k vyhodnocení extrakcí denotací. Pro extrakci denotací využívá autor neuronových sítí, což přináší velmi zajímavé výsledky. Řízení dialogu je navrženo nově a zajímavě, přičemž je prokázáno, že interakcí s uživatelem se system učí a může získat a dale využívat nové znalosti.

- Přestože autor nabízí teoretický princip fungování na otevřené doméně, zdá se, že implementovám a vyhodnocen byl pouze na jedné doméně – vyhledávání restaurací. Má disertant již praktické zkušenosti s vícedoménovými dialogy?
- Testování dialogového systému proběhlo na omezeném počtu konceptů a pouze text-based. Mohl by se disertant vyjádřit k otázce správnosti odvozených faktů v případě "očekávané" chybovosti ASR a NLU.
- Jak by se větší množství koncptů vyžadujících ohodnocení velkého množství paralelzních řetězců událostí podepsalo na schopnosti systému provozovat dialog v reálném čase?