

Dialogue Systems for Intelligent Human Computer Interactions

Erinc Merdivan¹

*AIT Austrian Institute of Technology GmbH, Wiener Neustadt, Austria
CentraleSupélec, Metz, France*

Deepika Singh²

*AIT Austrian Institute of Technology GmbH, Wiener Neustadt, Austria
Holzinger Group, HCI-KDD, Institute for Medical Informatics/Statistics, Medical University Graz,
Austria*

Sten Hanke³

AIT Austrian Institute of Technology GmbH, Wiener Neustadt, Austria

Andreas Holzinger⁴

*Holzinger Group, HCI-KDD, Institute for Medical Informatics/Statistics, Medical University Graz,
Austria*

Abstract

The most fundamental communication mechanism for interaction is dialogues involving speech, gesture, semantic and pragmatic knowledge. Various researches on dialogue management have been conducted focusing on standardized model for goal oriented applications using machine learning and deep learning models. The paper presents the overview on existing methods for dialogue manager training; their advantages and limitations. Furthermore, a new image-based method is used in Facebook bAbI Task 1 dataset in Out Of Vocabulary setting. The results show that using dialogue as an image performs well and helps dialogue manager in expanding out of vocabulary dialogue tasks in comparison to Memory Networks.

Keywords: dialogue system, image-based method, chatbots

¹ Email: erinc.merdivan@ait.ac.at

² Email: deepika.singh@ait.ac.at

³ Email: sten.hanke@ait.ac.at

⁴ Email: andreas.holzinger@medunigraz.at

1 Introduction

The Information and Communication Technology (ICT) with which we interact in daily life is more distributed and embodied into the environment (the so called intelligent space) [1]. Especially, when designing ICT solutions for elderly people, who are very often critical towards new technology, distributed system can be even more challenging. To improve the Human Computer Interaction (HCI) with ICT solutions, a directed natural interaction and an emotional intelligence is very important [2].

User studies regarding elder behavior change over the ageing process [3] identified that *"a skill that many elderly people retain, even with significant cognitive degradation, is the ability to communicate in a multimodal face-to-face fashion. The skills for this type of interaction are acquired in infancy and early childhood and comprise of tacit, crystallized knowledge in older adulthood"*. Face-to-face interaction incorporates a wide range of non-verbal and paraverbal ways to carry semantic content complementary to the speech. It allows persons with disabilities to compensate some perception channels (e.g., hearing) by utilizing other channels of communication (e.g., body gestures). Face-to-face dialog is also characterized by well-established repair mechanisms of understanding thus enabling the listener to request a repetition or clarification by the speaker (e.g., communicated by a head nod). Moreover, face-to-face dialog has built-in mechanisms for constraining the interactants focus of attention. This focus is important as some elderly people have difficulty dividing their attention or handling distractions.

This kind of face-to-face interaction can be provided utilizing avatars. Avatars have the potential to impersonate the used technology and thus increase the acceptance of the software [4]. The interaction with avatars is able to provide multiple advantages. Avatars can, for instance, provide gestures, which in turn are able to increase the understanding of the presented information. Furthermore, the visual enrichment of verbal information i.e., adding a lip synched animated character to audio speech output, can increase the intelligence and enhance the robustness of the information transmission as known from natural speech. Therefore, a consistency between the visual and vocal output is of uttermost important [5].

A growing amount of Embodied Conversational Agents (ECAs) make use of Natural Language Processing (NLP) for implementing the intelligent dialogue component. NLP-based systems are of major interest in human-machine interactions for multimodal interfaces and are preferred widely for Natural and Spoken Language Understanding (NLU or SLU), Natural Language Generation (NLG) and Dialogue Processing (DP) tasks.

Complete dialogue solution as shown in figure 1, consists of many parts, each of which specializes in certain task. Automatic Speech Recognition (ASR) module is responsible for converting spoken users utterance into text. Natural Language Interpreter converts textual information to meaningful features so that Dialogue State Tracker (DST) can process this features and update the current dialogue state. DST outputs current dialogue state so that Dialogue Response Selection(DRS) module

(which is trained to output a response to user utterance) can generate textual reply to user. Later this textual reply is converted to speech by text-to-speech (TTS) synthesizer. Since ASR and TTS are not related to dialogue manager directly, they can be considered as complementary modules to complete dialogue solution [14].

A dialogue manager is a component of a dialogue system, which is responsible for the transmission of information among participants in human-machine interaction. Dialogue manager can be divided into two main groups: Chatbots [16] and frame-based dialogues [17]. In chatbots, an agent is often trained to work without any knowledge of the dialogue structure mostly referred as open dialogue, whereas in frame-based dialogues; dialogue frames are structured by experts with slots and values that each slot can take. In this work, we mainly focus on chatbots and how they are trained with rule-based, sequence to sequence learning, reinforcement learning and hierarchical reinforcement learning based manager models. All these methods can be used separately or together to leverage each other's strengths. The paper gives an overview on the ECAs from the various projects; existing models applied to dialogue manager training and its advantages and drawbacks. Furthermore, it includes results on dialogue dataset with new image based method where dialogue is processed as images to train dialogue manager.

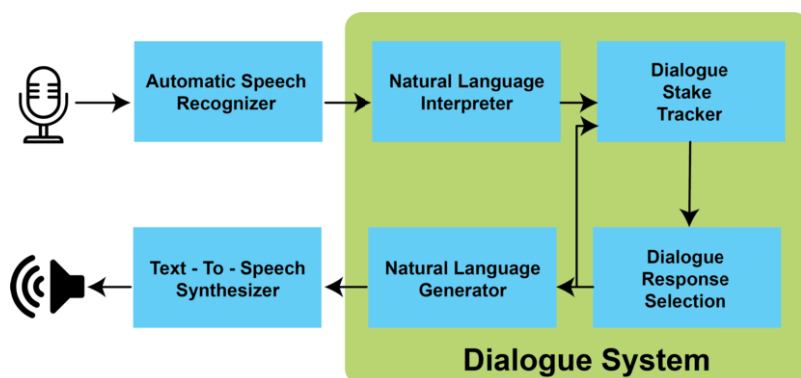


Fig. 1. Schematic diagram of Complete Dialogue System

The paper is structured into six sections. First section provides an introduction of the Dialogue system followed by Embodied Interaction. The third section describes the existing rule based and machine learning based models in dialogue systems. Later, a new proposed methodology for the dialogue system is explained. The next section consists of the preliminary results and their discussions; and the last section concludes the findings of the work.

2 Embodied Interaction Experience

Several working groups and projects have been shown that embodied interactions can help to increase the interaction with ICT systems [1]. Bickmore et al. developed a virtual laboratory to explore the longitudinal usage of a virtual exercise coach. Older adult participants interacted with an agent from their home once a day for up to 120 days. The results showed that users who interacted with an

agent that used variable dialogue exercised significantly more than those interacting with an ECA with non-variable dialogue. It has been shown that the natural dialogue is an important part in natural conversation, specifically how an ECA is communicating and is structuring the dialogue with the user. In this way ECAs are capable of understanding the user's voice, gestures and emotions and plan based on this understanding multimodal utterances from propositional representations of meaning [8]. In Embodied Interactions an intelligent dialogue manager component is responsible for interpreting the users speech commands and generating the appropriate verbal and non-verbal behaviour of the agent. Studies performed in our projects: Miraculous-Life⁵ [12], CaMeLi⁶ [11] and Ibi [13] have shown that the embodiment (the avatar) should express appropriate empathetic feedback (i.e., emotional facial expressions and voice intonation) based on behaviour and dialogue patterns established between older adults and human caregivers over the lifetime. In terms of dialogue communication, the avatar should allow smooth turn-taking by different modes like waiting, listening and talking. Figure 2 shows a screenshot of the Miraculous-Life animated avatar. In terms of interaction, the results indicated that the communication of older adults with the agent should be based on friendly and natural dialogues and it is critical that the agent has a clearly understandable speech output [9] and [10]. As some older adults experienced difficulties hearing and understanding the agents' speech, pronunciation and intonation has to be improved and more preferences are given to the agent which can adapt to the native languages of the user [15]. It has been observed that interaction with ECAs is still not entirely multimodal, rather more speech command based, which is also a cause of the limited flexibility of existing dialogue components. Furthermore, the most reported reason was that the ECAs did not fully reach users original expectations. Among the main reasons for this we note the mismatched expectations related with verbal communication capabilities. In fact, the majority of the users in the target group, on facing a human-like character expected a more natural interaction in terms of speech dialogues. This includes the Automatic Natural Speech Recognition but also the flexibility of the dialogue itself. Participants easily got frustrated after a few unexpected verbal behaviour by the agent. On the other hand, we noticed that users from the target groups face some challenges regarding training the interaction with the agent. This led to a higher number of repetitions than desirable. From these consideration, two main conclusions have been drawn.

- (i) An ideal solution would be to have more flexibility and variety in the speech commands.
- (ii) it is of utmost importance that all the interaction components run as robustly as possible, are fault-tolerant, and support repair mechanisms.

⁵ <http://www.miraculous-life.eu/>

⁶ <http://www.aal-europe.eu/projects/cameli/>

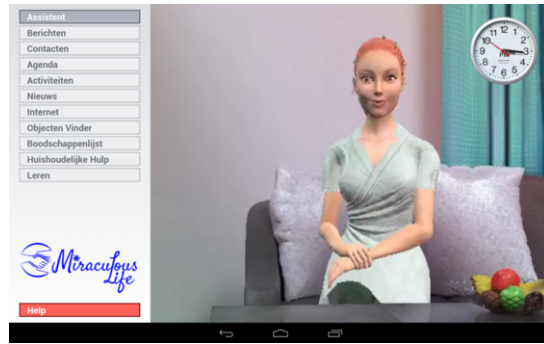


Fig. 2. The Miraculous-Life user interface including the animated avatar

3 The existing dialogue systems

The section provides the overview of the existing work on training dialogue managers which includes: Rule-based methods; Sequence-to-Sequence methods; Reinforcement learning based methods; and Hierarchical reinforcement learning methods.

3.1 Rule-based methods

The first chatbot developed by using rule-based system was ELIZA [18], which uses pattern matching based on user replies. In rule-based systems, human dialogues are modeled as set of states and dialogue manager has to choose replies for the conversation from the given set of rules [19]. This model has been used in many different applications such as restaurant booking or online psychological therapy chatbot [18]. In this method, a human, who is usually a domain expert analyzes the dialogue flow between human agents and tries to come up with predefined dialogue states and possible replies for each state based on patterns. The advantages of rule-based methods are that dialogue managers have control on selecting replies for the conversation and these selected replies from the full set of replies ensure that the user is not upset or offended, thus keeping the system consistent.

is a dialogue response [23,24]. In translation tasks, dataset consists of pair of sentences in different languages, whereas in the dialogue tasks, there can be many more dialogue replies for the same history.

Although, seq2seq suffers from one to many mapping which has been applied to train dialogue management tasks. It has given good results on chatbots and also in domain specific tasks such as restaurant booking. Since these methods give most probable answer, it is often difficult to customize unless additional information is given and may overfit in generating certain frequent replies in the dataset. A solution could be by adding embeddings about user information in order to train network to take user preferences into account. [25].

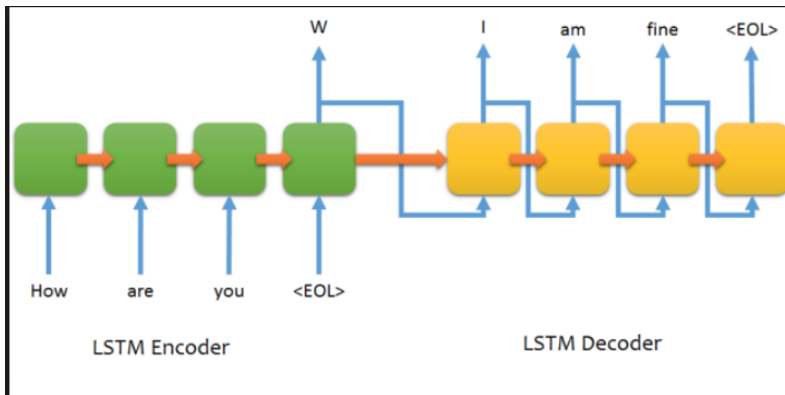


Fig. 4. Sequence-to-Sequence learning [22]

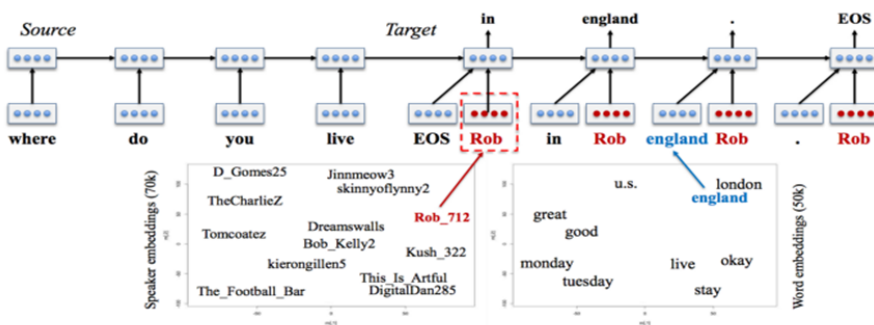


Fig. 5. Persona based neural response generation [25]

The figure 4 demonstrate seq2seq learning applied on generating response for the given history. LSTM encoder takes either full history or last reply and converts it into encoded feature vector where LSTM decoder takes this vector and outputs a possible reply condition on encoded feature vector. In figure 5, seq2seq model is changed to also accept and embedding of user, 'Rob' in this case, which allows network to generate replies based on user persona.

3.3 Reinforcement learning based methods

Reinforcement learning (RL) method is a machine learning method where agent learns to maximize notion of reward in given environment by learning which actions to take in each state [34]. Reinforcement learning methods gained popularity in recent years due to revisiting existing RL algorithms with new deep learning methods. Deep reinforcement learning algorithms outperform humans in certain game tasks such as in Atari games, chess and GO [27].

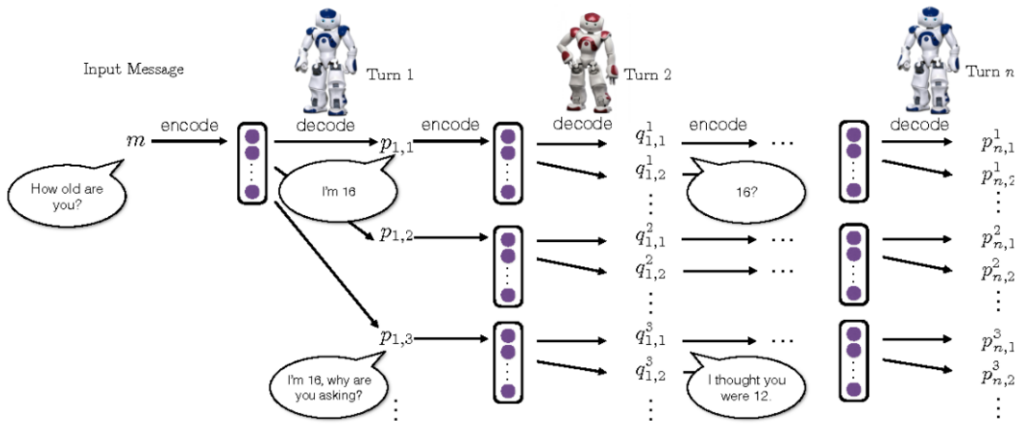


Fig. 6. Dialogue simulation between agents [29]

Dialogue management can be formulated as a Markov Decision Process which is defined as a tuple (S, A, T, R, γ) . S and A represent set of states and actions; T represent the transition probabilities between s and s' given action a is taken in state s . R is the reward function such as $R: S \times A \rightarrow \mathbb{R}$ which gives the reward value for an action a taken in state s . γ is the discount factor which is used to calculate the expected discounted cumulative reward $\mathbb{E}[\sum_{i \geq 1} \gamma^{i-1} R_{t+i}]$. policy π of RL agent is

a mapping between states and actions. Policy of an agent can either be stochastic or deterministic and actions policy maps to can be either continuous valued actions or discrete valued actions. RL agent dialogue manager is trained in order to learn a policy which maximizes the expected discounted cumulative return.

Deep RL algorithms can be even trained playing against each other, to achieve post human performance even when the state space is very high. RL methods require a scalar reward which is more natural for dialogue tasks than supervised models. Since, it is difficult to get a supervised dialogue dataset; it would be more natural for RL dialogue manager to discover itself the dynamics of dialogue by following a reward function. However, in dialogue tasks it is often hard to design such a reward for RL methods to trained on dialogue tasks [28]. Heuristic rewards can be designed by human experts in order to train dialogue manager with RL based methods [29] or automated dialogue reply measures can also be used for training [30] [31]. The figure 6 shows a reinforcement learning framework for neural response generation by simulating dialogues between the two agents, integrating the

strengths of neural seq2seq systems and reinforcement learning for dialogue.

In game setting tasks, it is also important to note that action space is very low whereas in dialogue tasks they are high dimensional. This high dimensionality is due to large number of possible sentences in any language. Such high number of possible actions increases the complexity of RL methods to apply on dialogue tasks. Often less importance is given to number of training samples needed for RL methods. For even simple data efficient methods it may require 10K of training data which is impractical to collect from humans and also impractical to simulate due to complexity and variety in human dialogues [32].

3.4 Hierarchical Reinforcement learning-based methods

Dialogue modelling can be modeled as an Markov Decision Process (MDP), however, if dialogue task is complex and natural language is used, then state and action spaces are of very high dimension. Traditional RL suffers from curse of dimensionality and different solutions are investigated by researchers. One of solutions, is using hierarchical reinforcement learning, where agent learns to abstract problem. Agent in HRL setting learns to abstract state and/or action instead of learning primitive actions for each state. The agent usually has a different hierarchical levels of policies where high level policy controls which low level policy would be chosen; while each low level policy is optimized for a different and simpler task. Furthermore, the agent makes a decision on which policy it should choose from and then follow this sub-policy which is different from higher level policy until termination condition is satisfied [33].

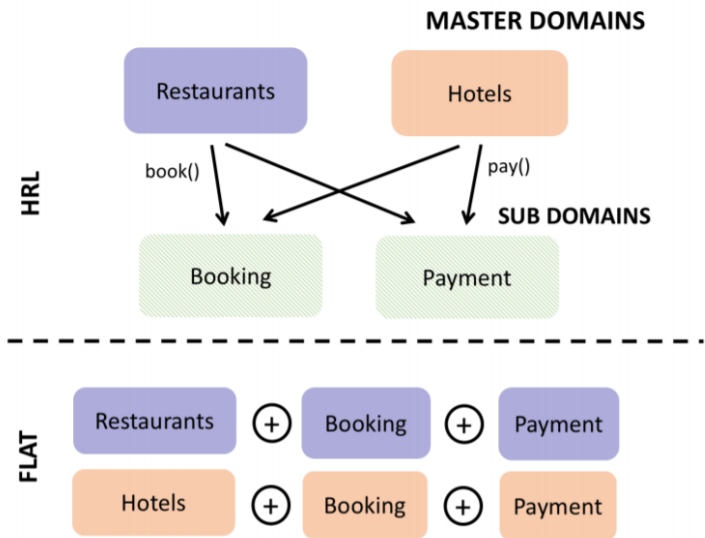


Fig. 7. Comparison of two analyzed architecture [36]

HRLs are mainly applied on special version of MDP call Semi-MDPs which allows actions to last for different amount of time between state transitions in order

to model temporally-extended action set [34]. First evaluation of HRL methods on SMDP-based dialogue agent setting on a realistic environment, showed that semi-learned or hierarchical behaviours outperformed fully-learned agents behaviour in a travel planning domain with real users [35]. Difference between flat(traditional RL) and hierarchical RL for Booking and Payment tasks/sub-domains shared between two master domains is shown in figure 7, where flat RL learns only about primitive actions and needs to output an action for each time step whereas HRL learns hierarchical abstraction and may learn primitive as well as composite actions.

In more recent work, [36] introduced hierarchical model which is trained by hierarchical reinforcement learning with the Gaussian Process as function approximators; while [37] uses deep Q-networks as function approximators and both HRL models outperform standard RL models over composite tasks. HRL based methods on dialogue management although outperforms standard RL, have open research questions concerning handling deeper hierarchies; designing reward for lower level hierarchies; and also automatically dividing complex task into simpler subtasks which is possible with deeper hierarchies than few levels of hierarchy.

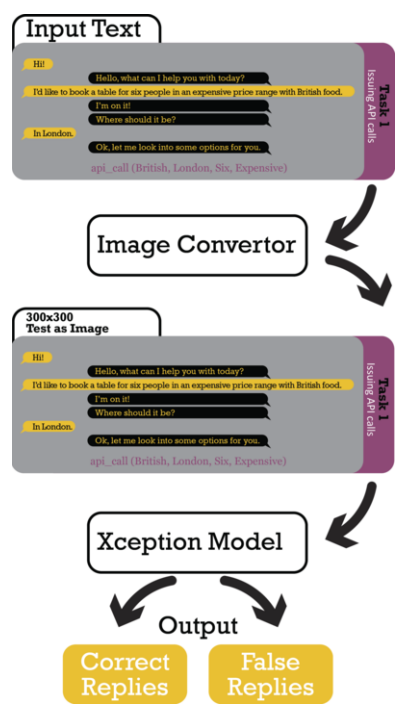


Fig. 8. Image-based methodology for dialogue manager

4 Image-based method

Recently, a novel method has been proposed to train dialogue managers [38]. In this method dialogue is converted to an image by rendering the dialogue history with an reply as an image. Processing sequential data as an image is quite common in audio processing of spectrograms. Although, audio is sequential and recurrent neural

networks or 1-D convolutional neural networks are more suited, hence, they can also be processed as an image. Significant research works have been conducted using spectrograms and implementation of vision based methods in audio domain [39–41].

In dialogue domain (figure 8), same idea has been applied where manager is trained by creating positive samples combining dialogue history with correct reply and negative samples combining dialogue history with false replies samples from dataset of possible replies. As shown in figure 9, each dialogue in this image are 300x300 black and white images and processed as an image using Xception vision model [42]. The above two blocks of the figure consists of correct dialogues where dialogue histories are appended with correct replies, given as an image. And, in lower two blocks, the same dialogue histories are appended with false replies, which are sampled from the all possible candidate replies. It can be seen that dialogue manager task is to provide correct information to the user, who asked about restaurant address. In above two images (dialogue with correct replies) dialogue manager returns address of correct restaurant, however, in lower two images (dialogue with false replies) dialogue manager replies with irrelevant replies such as offering new restaurant on left and making api call on right down. In such a setup, dialogue manager investigates dialogue only visually and learns which dialogues are correct and which dialogues are false. However, this methodology has never been applied in dialogue modeling.

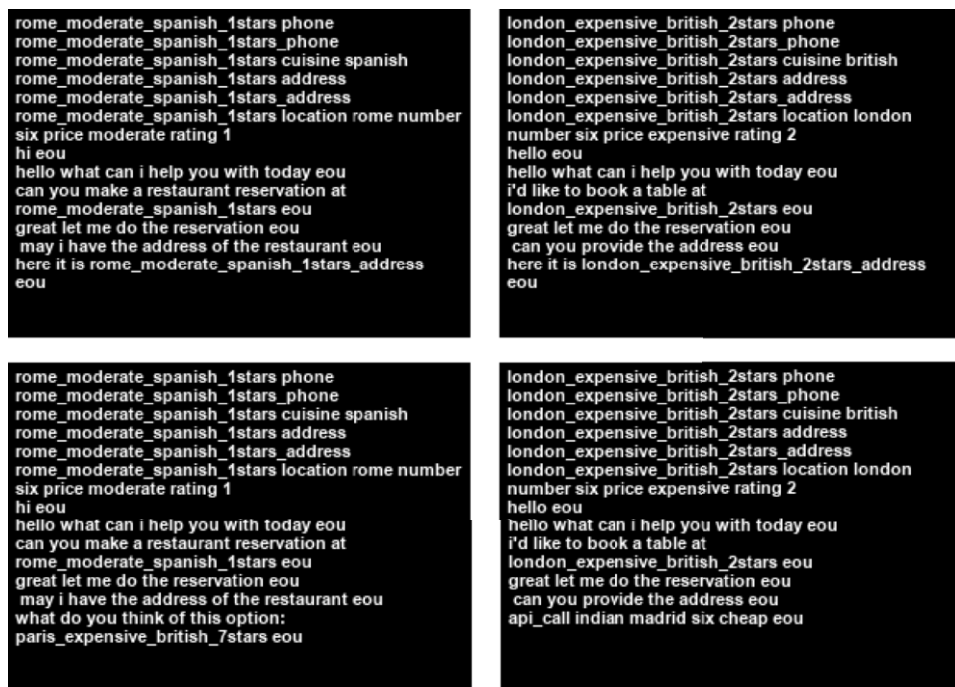


Fig. 9. Dialogue as an image

5 Experiment and Results

We used Xception vision model on Facebook bAbI dataset Task 1 [21]. In this approach, we created a train set consisting of correct and wrong dialogue history and reply pairs. Later, we converted this textual dialogue to an image and processed it as a binary classification problem. For testing we rank all candidate answers for given dialogue history and choose the highest. This work is an extension of the work done [38] in Out Of Vocabulary (OOV) setting.

In natural language domain dictionary, it is important when manager does not have certain words in its dictionary than understanding the text is not possible. MemNets performance drops when tested on OOV, since language based models fails when they do not find OOV words during training. Adding human generated features can overcome this issue, however, it is very task dependent and requires human effort. On the other hand, it is different in image domain situations since there is no need for dictionary instead the network should only learn to match patterns of images. Nevertheless, other issues such as font type and font size may arise in image domain with image based methods. However, it is out of the scope of this work and therefore it is not required to have a method to adopt different font types and sizes since the aim is to learn dialogue flow.

In Facebook bAbI dataset Task 1 there are 1000 training and 1000 test dialogues. We sampled 10 negative responses for each training dialogue and we chose the correct answer from 4212 candidates answers for test dialogues. Experiments are done on a single GPU (1080ti) and took a week’s time to train the model. As reported in Table 1, the results show that using image domain also helps the dialogue manager to expand to Out Of Vocabulary (OOV) dialogue tasks naturally where OOV response accuracy matches in vocabulary results.

Table 1
Facebook bAbI Dialog Task 1

Metrics	Xception (%)	Memory Networks w/o Match Type (%)	Memory Networks w/ Match Type (%)
Per-response Accuracy	85.7	99.9	100.0
Per-response Accuracy (Out-Of-Vocabulary)	86.6	72.3	100.0

6 Conclusion

A new image-based approach to train dialogue managers has been implemented on Facebook bAbI dialog task 1. Our method outperforms the Memory Networks (without any extra processing such as match type) for dialogue response accuracy in OOV dialog tasks using 10% of training dataset. Performance is consistent between in vocabulary and OOV. Furthermore, the experiments on English and other languages with large dataset are needed to fully capture the benefits of image based processing. It will also be interesting to perform both image-based and language-based methods simultaneously. Further work will be performed by implementing

the improved dialogue components and testing it with other embodied interaction components.

Acknowledgement

This work has been funded by the European Union Horizon2020 MSCA ITN ACROSSING project (GA no. 616757). The authors would like to thank the members of the project's consortium for their valuable inputs.

References

- [1] Sten Hanke, Christiana Tsiourti, Miroslav Sili, Eleni Christodoulou, *Embodied Ambient Intelligent Systems*, Ambient Intelligence and Smart Environments, Volume **20**: Recent Advances in Ambient Assisted Living Bridging Assistive Technologies, e-Health and Personalized Health Care, 65–85, (2015)
- [2] P.H. Robert, A. König, H. Amieva, S. Andrieu, F. Bremond, R. Bullock, M. Ceccaldi, B. Dubois, S. Gauthier, P.-A. Kenigsberg, S. Nave, J.M. Orgogozo, J. Piano, M. Benoit, J. Touchon, B. Vellas, J. Yesavage and V. Manera, *Recommendations for the use of Serious Games in people with Alzheimers Disease*, related disorders and frailty, in Front Aging Neurosci., 2. Mrz (2014)
- [3] T. W. Bickmore, L. Caruso, E. Clough-Gorr, *Acceptance and usability of a relational agent interface by urban older adults*, In CHI05 extended abstracts on Human factors in computing systems, ACM, Press, Portland, OR, USA, 1212-1215, (2005)
- [4] A. H. Martin M. Morandell, B. Wöckl, S. Dittenberger und S. Fagel, *Avatars@Home: Interfacing the Smart Home for Elderly People*, in HCI and Usability for e-Inclusion: 5th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2009, Linz, Austria, November 9-10, 2009 Proceedings, Linz, Austria, Springer Berlin Heidelberg, 353-365, (2009)
- [5] S. Fagel, A. Hilbert, C. Mayer, M. Morandell, M. Gira und M. Petzold, *Avatar User Interfaces in an OSGi-based System for Health Care Services*, Global Health 2013, The Second International Conference on Global Health Challenges, 1-4, 17 November (2013).
- [6] R. Sharma, M. Yeasin, N. Krahtoeffer, I. Rauschert, I. Brewer, A.M. Maceachren and K. Sengupta, *Speech-gesture driven multimodal interfaces for crisis management*, Proceedings of the IEEE, Volume 91(9), p. 1327-1354, (2003).
- [7] T.H. Bui, *Multimodal Dialogue Management - State of the art*, Centre for Telematics and Information Technology University of Twente. Retrieved from <http://eprints.ewi.utwente.nl/5708/>, 2006
- [8] S. Kopp, L. Gesellensetter, N. C. Krämer, I. Wachsmuth, *A Conversational Agent as Museum Guide Design and Evaluation of a Real-World Application Intelligent Virtual Agents*, Lecture Notes in Computer Science Volume **3661**, p. 329-343, (2005).
- [9] M. Sili, J. Bobeth, E. Sandner, S. Hanke, S. Schwarz, C. Mayer, *Talking Faces in Lab and Field Trials - A View on Evaluation Settings and User Involvement Results of Avatar Based User Interaction Techniques in Three Ambient Assisted Living Projects*, International Conference on Human Aspects of IT for the Aged Population, pp. 134-144, Springer, Cham, (2015).
- [10] D. Cereghetti, C. Wings, J. Meijers, D6.4 Pilot acceptance evaluation results, public deliverable, <http://www.miraculous-life.eu/public-deliverables> last seen: 10.11.2018
- [11] Christiana Tsiourti, Joo Quintas, Maher Ben-Moussa, Sten Hanke, Niels Alexander Nijdam, Dimitri Konstantas, *The CaMeLi Framework A Multimodal Virtual Companion for Older Adults*, Proceedings of SAI Intelligent Systems Conference, pp. 196-217, Springer, Cham, 2016/9/21
- [12] Sten Hanke, Emanuel Sandner, Samat Kadyrov, Andreas Stainer-Hochgatterer, *Daily life support at home through a virtual support partner*, IET Digital Library, 2016/1/1
- [13] Miroslav Sili, Jan Bobeth, Emanuel Sandner, Sten Hanke, Stephanie Schwarz, Christopher C. Mayer, *Talking Faces in Lab and Field Trials - A View on Evaluation Settings and User Involvement Results of Avatar Based User Interaction Techniques in Three Ambient Assisted Living Projects.*, pp. 134-144, HCI (25) 2015
- [14] Serban, Iulian Vlad, Ryan Lowe, Peter Henderson, Laurent Charlin, and Joelle Pineau, *A survey of available corpora for building data-driven dialogue systems* URL:<https://arxiv.org/pdf/1512.05742.pdf> (2015).

- [15] Singh, Deepika, Ismini Psychoula, Johannes Kropf, Sten Hanke, and Andreas Holzinger, *Users Perceptions and Attitudes Towards Smart Home Technologies*, In International Conference on Smart Homes and Health Telematics, 203-214. Springer, Cham, (2018).
- [16] Shawar, Bayan Abu, and Eric Atwell, *Chatbots: are they really useful?*, In Ldv Forum, vol. **22**, no. **1**, (2007), 29-49.
- [17] Masche, Julia, and Nguyen-Thinh Le, *A Review of Technologies for Conversational Systems*, International Conference on Computer Science, Applied Mathematics and Applications. Springer, Cham, (2017).
- [18] Weizenbaum, Joseph, *ELIZA a computer program for the study of natural language communication between man and machine* Communications of the ACM **9.1** (1966): 36-45.
- [19] Webb, Nick, *Rule-based dialogue management systems* In Proceedings, (2000), 164-169.
- [20] Berg, Markus M., *NADIA: A Simplified Approach Towards the Development of Natural Dialogue Systems*, In International Conference on Applications of Natural Language to Information Systems, pp. 144-150. Springer, Cham, (2015).
- [21] Bordes, Antoine, Y-Lan Boureau, and Jason Weston *Learning end-to-end goal-oriented dialog*, URL: <https://arxiv.org/abs/1605.07683>
- [22] Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le, *Sequence to sequence learning with neural networks* Advances in neural information processing systems, (2014).
- [23] Vinyals, Oriol, and Quoc Le, *A neural conversational model*, URL: <https://arxiv.org/pdf/1506.05869.pdf> (2015).
- [24] Sordoni, Alessandro, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Margaret Mitchell, Jian-Yun Nie, Jianfeng Gao, and Bill Dolan, *A neural network approach to context-sensitive generation of conversational responses*, URL: <https://arxiv.org/pdf/1506.06714.pdf> (2015)
- [25] Li, Jiwei, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, and Bill Dolan, *A persona-based neural conversation model*, URL: <https://arxiv.org/pdf/1603.06155.pdf> (2016)
- [26] Sutton, Richard S., and Andrew G. Barto, *Reinforcement learning: An introduction* (2011).
- [27] Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves et al, *Human-level control through deep reinforcement learning* Nature 518, no. **7540** (2015), 529.
- [28] Liu, Chia-Wei, Ryan Lowe, Iulian V. Serban, Michael Noseworthy, Laurent Charlin, and Joelle Pineau, *How not to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation*, URL: <https://arxiv.org/pdf/1603.08023.pdf> (2016).
- [29] Li, Jiwei, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, and Dan Jurafsky, *Deep reinforcement learning for dialogue generation*, URL: <https://arxiv.org/pdf/1606.01541.pdf> (2016).
- [30] Merdivan, Erinc, Mohammad Reza Loghmani, and Matthieu Geist, *Reconstruct and Crush Network*, Advances in Neural Information Processing Systems, (2017).
- [31] Lowe, Ryan, Michael Noseworthy, Iulian V. Serban, Nicolas Angelard-Gontier, Yoshua Bengio, and Joelle Pineau, *Towards an automatic Turing test: Learning to evaluate dialogue responses*, URL: <https://arxiv.org/pdf/1708.07149> (2017).
- [32] Pietquin, Olivier, Matthieu Geist, Senthilkumar Chandramohan, and Herv Frezza-Buet, *Sample-efficient batch reinforcement learning for dialogue management optimization* ACM Transactions on Speech and Language Processing (TSLP) **7**, no. **3** (2011), 7.
- [33] Barto, Andrew G., and Sridhar Mahadevan, *Recent advances in hierarchical reinforcement learning*, Discrete event dynamic systems **13** **1-2** (2003), 41-77.
- [34] Sutton, Richard S., Doina Precup, and Satinder Singh. *Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning* Artificial intelligence **112**, no. **1-2** (1999), 181-211.
- [35] Cuayhuatl, Heriberto, Steve Renals, Oliver Lemon, and Hiroshi Shimodaira, *Evaluation of a hierarchical reinforcement learning spoken dialogue system* Computer Speech and Language **24**, no. **2** (2010), 395-429.
- [36] Budzianowski, Pawe, Stefan Ultes, Pei-Hao Su, Nikola Mrki, Tsung-Hsien Wen, Inigo Casanueva, Lina Rojas-Barahona, and Milica Gai, *Sub-domain modelling for dialogue management with hierarchical reinforcement learning* URL: <https://arxiv.org/pdf/1706.06210.pdf>, (2017).

- [37] Peng, Baolin, Xiujun Li, Lihong Li, Jianfeng Gao, Asli Celikyilmaz, Sungjin Lee, and Kam-Fai Wong, *Composite task-completion dialogue policy learning via hierarchical deep reinforcement learning*, URL: <https://arxiv.org/pdf/1704.03084.pdf>, (2017).
- [38] Merdivan, Erinc, Anastasios Vafeiadis, Dimitrios Kalatzis, Sten Henke, Johannes Kropf, Konstantinos Votis, Dimitrios Giakoumis et al, *Image-based Natural Language Understanding Using 2D Convolutional Neural Networks*, URL: <https://arxiv.org/pdf/1810.10401.pdf>, (2018).
- [39] Abdel-Hamid, Ossama, Abdel-rahman Mohamed, Hui Jiang, Li Deng, Gerald Penn, and Dong Yu, *Convolutional neural networks for speech recognition*, IEEE/ACM Transactions on audio, speech, and language processing 22, no. 10 (2014), 1533-1545.
- [40] Phan, Huy, Lars Hertel, Marco Maass, and Alfred Mertins, *Robust audio event recognition with 1-max pooling convolutional neural networks*, URL: <https://arxiv.org/pdf/1604.06338.pdf>, (2016).
- [41] Hershey, Shawn, Sourish Chaudhuri, Daniel PW Ellis, Jort F. Gemmeke, Aren Jansen, R. Channing Moore, Manoj Plakal et al, *CNN architectures for large-scale audio classification* In Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on, 131-135. IEEE, (2017).
- [42] Chollet, Franois, *Xception: Deep learning with depthwise separable convolutions*, URL: <https://arxiv.org/pdf/1610.02357.pdf>, (2017).