Dialogue Systems Using Online Learning: Beyond Empirical Methods *

Heriberto Cuayáhuitl
German Research Center for Artificial Intelligence
Saarbrücken, Germany
`hecu01@dfki.de`

Nina Dethlefs
Heriot-Watt University
Edinburgh, Scotland
`n.s.dethlefs@hw.ac.uk`

Abstract

We discuss a change of perspective for training dialogue systems, which requires a shift from traditional empirical methods to online learning methods. We motivate the application of online learning, which provides the benefit of improving the system's behaviour continuously often after each turn or dialogue rather than after hundreds of dialogues. We describe the requirements and advances for dialogue systems with online learning, and speculate on the future of these kinds of systems.

1 Motivation

Important progress has been made in empirical methods for training spoken or multimodal dialogue systems over the last decade. Nevertheless, a different perspective has to be embraced if we want dialogue systems to learn on the spot while interacting with real users. Typically, empirical methods operate cyclically as follows: collect data, provide the corresponding annotations, train a statistical or other machine learning model, evaluate the performance of the learned model, and if satisfactory, deploy the trained model in a working system. The disadvantage of this approach is that while data is still being collected subsequent to deployment, the system does not optimize its behaviour anymore (cf. stepwise learning, the solid blue line in Fig. 1). In contrast, dialogue systems with online learning tackle this limitation by learning a machine learning model
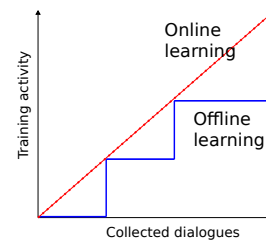
Figure 1: Learning approaches for dialogue systems. Whilst offline learning aims for discontinuous learning, online learning aims for continuous learning while interacting with users in a real environment.

continuously often from unlabeled or minimally labeled data (cf. dotted red line in Fig. 1). So whilst empirical methods train models after hundreds of dialogues, online learning methods refine the system models after each user turn or each dialogue. In the rest of the paper we discuss the requirements, advances and potential future of these kind of systems.

2 Online Learning Systems: Requirements

Several requirements arise for the development of successful online learning systems. First of all, they need to employ methods that are **scalable** for real-world systems and the modelling of knowledge in sufficient detail. Second, **efficient learning** is a prerequisite for learning from an ongoing interaction without causing hesitations or pauses for the user. Third, learnt models should satisfy a **stability** criterion that guarantees that the learning agent's performance does not deteriorate over time, e.g. over the course of a number of interactions, due to the newly accumulated knowledge and behaviours. Fourth,

7

systems should employ a **knowledge transfer** approach in which they master new tasks they are confronted with over their life span by transferring general knowledge gathered in previous tasks. Fifth, online learning sytems should adopt a **lifelong learning** approach, arguably without stopping learning. This implies making use of large data sets, which can be unlabeled or partially labeled due to the costs that they imply. Finally, in the limit of updating the learned models after every user turn, the online and offline learning methods could be the same as long as they meet the first three requirements above.

3   Online Learning Systems: Advances

Several authors have recognised the potential benefits of online learning methods in previous work.

Thrun (1994) presents a robot for lifelong learning that learns to navigate in an unknown office environment by suggesting to transfer general purpose knowledge across tasks. Bohus et al. (2006) describe a spoken dialogue system that learns to optimise its non-understanding recovery strategies online through interactions with human users based on pre-trained logistic regression models. Cuayáhuitl and Dethlefs (2011) present a dialogue system in the navigation domain that is based on hierarchical reinforcement learning and Bayesian Networks and relearns its behaviour after each user turn, using indirect feedback from the user's performance. Gašić et al. (2011) present a spoken dialogue system based on Gaussian Process-based Reinforcement Learning. It learns directly from binary feedback that users assign explicitly as rewards at the end of each dialogue and that indicate whether users were happy or unhappy with the system's performance. From these previous investigations, we can observe that online learning systems can take both explicit and/or implicit feedback to refine their trained models.

4   Online Learning Systems: Future

While previous work has made important steps, the problem of lifelong learning for spoken dialogue systems is far from solved. Especially the following challenges will need to receive attention: (a) fast learning algorithms that can retrain behaviours after each user turn with stable performance; and (b) scalable methods for optimizing multitasked behaviours at different levels and modalities of communication.

In addition, we envision online learning systems with the capability of transfering knowledge across systems and domains. For example: a dialogue act classifier, an interaction strategy, or a generation strategy can be made transferable to similar tasks. This could involve reasoning mechanisms to infer what is known/unknown based on past experiences. The idea of learning from scratch every time a new system is constructed will thus be avoided. In this regard, the role of the system developer in these kinds of systems is not only to specify the system's tasks and learning environment, but to constrain and bootstrap the system behaviour for faster learning. All of these capabilities will be possible using online learning with a lifelong learning perspective.

5   Tools and Data

Currently there are software tools for training models but they are more suitable for offline learning.[1] Software tools for online learning remain to be developed and shared with the community. In addition, since building a dialogue system typically requires a tremendous amount of effort, researchers working on learning approaches should agree on standards to facilitate system development. Finally, since dialogue data is an often lacking resource in the community, the online learning perspective may contribute towards reducing the typical chicken and egg problem, due to dialogue knowledge being more readily transferable across domains, subject to online adaption towards particular domains.

References

Dan Bohus, Brian Langner, Antoine Raux, Alan Black, Maxine Eskenazi, and Alexander Rudnicky. 2006. Online Supervised Learning of Non-Understanding Recovery Policies. In *Proc. IEEE SLT*.

Heriberto Cuayáhuitl and Nina Dethlefs. 2011. Optimizing Situated Dialogue Management in Unknown Environments. In *Proc. INTERSPEECH*.

Milica Gašić, Filip Jurčíček, Blaise Thomson, Kai Yu, and Steve Young. 2011. On-line policy optimisation of spoken dialogue systems via interaction with human subjects. In *Proc. IEEE ASRU*.

Sebastian Thrun. 1994. A Lifelong Learning Perspective for Mobile Robot Control. In *Proc. IEEE/RSJ/GI*.

[1]www.cs.waikato.ac.nz/ml/weka/