Language Technologies Institute

Thesis Proposal

GHC 4405| Tuesday, December 19th| 12:00 pm

Learning Generative End-to-end Dialog Systems with Knowledge

Tiancheng Zhao

COMMITTEE:

Maxine Eskenazi, (chair)

William Cohen

Louis-Philippe Morency

Dilek Hakkani – Tur (Google)

Abstract

Traditional dialog systems follow a modular approach and often have trouble expanding to new or more complex domains. The goal of this dissertation is to create domain-agnostic learning algorithms and dialog models that can continuously learn to converse in any new domains and only requires a small amount of new data for the adaption. Achieving this goal first requires powerful statistic models that are expressive enough to model the natural language and decision-making process in dialogs of many domains; second requires a learning framework enabling models to share knowledge from previous experience so it can learn to converse in new domains with limited data.

End-to-end (E2E) generative dialog models based on encoder-decoder neural networks are strong candidates for the first requirement. These models are not restricted to hand-crafted intermediate states and can in principle generalize to novel system responses. However, it is far from trivial to build a full-fledged dialog system using encoder-decoder models. Thus in the first stage of this thesis, we develop a set of novel neural network architectures that offer key properties that are required by dialog modeling. We tackle the second requirement by proposing a learning with knowledge (LWK) framework that can adapt the proposed system to new domains with minimum data. Two types of knowledge are studied: 1) domain knowledge from human experts 2) models' knowledge from learning related domains. We develop novel domain-aware models and training algorithms to teach the system to learn from data in related domains and generalize to unseen ones.

In conclusion, this dissertation shows that by combing specialized encoder-decoder models with the proposed LWK framework, E2E generative dialog models can be readily applied in complex dialog applications and expanded to new domains with limited resources, which we believe is an important step towards future conversational agents.

http://www.cs.cmu.edu/~tianchez/data/TianchezPhdProposal.pdf