Dialog System Learning New Concepts

Goal: Learn a dialog policy that trades off:
- Clarification to get necessary information for current user interaction
- Collecting labelled examples to build classifiers for new concepts
- Task Completion

We augment task oriented dialog systems with opportunistic active learning queries to enable the system to generalize to new concepts not seen at training time. For a dialog system in the shopping domain, this means automatically adapting to an expanded inventory with new product categories without collecting training dialogs on these categories.

Visual Attribute Classifier

Two branch convolutional neural network to upweight positive examples

Policy Learning:
- Dialog modeled as MDP with featureized state-action pairs
- Reward: Large positive value for correct guess, large negative value for incorrect guess, small negative value for additional questions
- RL algorithms: Q-learning and A3C

Results

Neither the static nor the learned policies transfer well during human evaluation but the learned policy remains more successful than the static policy.

Utility of Clarifications

Only the combination of learned clarifications and learned active learning queries improve over direct retrieval.

Experiment Phases:
- Classifier Initialization - Train classifier using paired images and labels
- Policy Initialization - Collect experience using the baseline to initialize the policy
- Policy Training - Improve the policy from on-policy experience
- Policy Testing - Policy weights are fixed, and we run a new set of interactions, reset classifiers to the state at the end of classifier initialization, over an independent test set containing novel attributes.

Policy Features:
- Clarification Policy Features - Metrics about current beliefs, information gain estimated from classifier probabilities
- Active Learning Policy Features - Margin, Fraction of previous uses and successes
- Decision Policy Features - Metrics about current beliefs, information gain, margin, dialog length

Static Baseline:
- Clarification: Choose query with maximum estimated information gain
- Active Learning: Uncertainty Sampling
- Decision Policy
  - Fixed dialog length
  - Dialog split equally between clarification and active learning
  - Heuristic checks to ensure usefulness of queries

Dialog Policy Learning for Joint Clarification and Active Learning Queries

Aishwarya Padmakumar¹
Amazon Alexa AI

Raymond J. Mooney²
University of Texas at Austin