Spoken Dialog Management with Partially Observable Markov Decision Processes

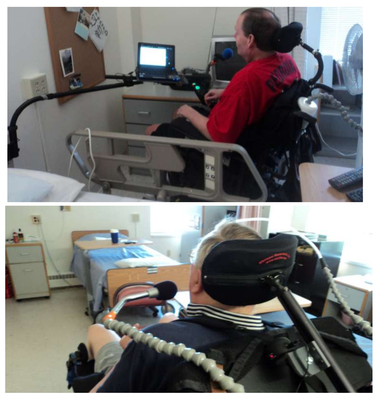
William Li
Advisors: Nicholas Roy, Seth Teller, James Glass

Motivation

- Speech-based assistive technology for adults with limited mobility and fine-motor skills at The Boston Home (TBH)
- High automatic speech recognition (ASR) error rates due to fatigue, dysphonia, and dysarthria
- Approach: Use **dialog** (requests for repetition, confirmation, or clarification) to infer user goals

ASR Concept Error Rates

Lab Speakers	Error Rate	TBH Speakers	Error Rate
lab01	4.0%	tbh01	12.0%
lab02	7.4%	tbh02	3.7%
lab03	10.9%	tbh03	5.1%
lab04	4.3%	tbh04	34.6%
lab05	12.7%	tbh05	57.1%
lab06	3.3%	tbh06	26.1%
lab07	3.8%	tbh07	9.4%
mean	**5.7%**	**mean**	**25.1%**
std. dev.	**4.5%**	**std. dev.**	**19.5%**



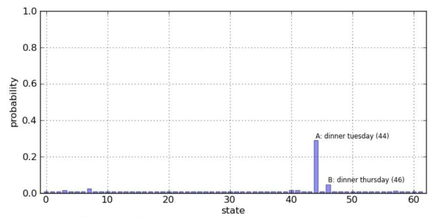
Bedside (top) and wheelchair (bottom) speech systems

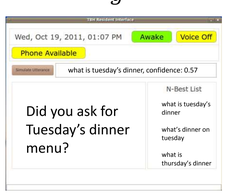
Spoken Dialog System POMDP Formulation

Entity	Description	Example
States, *G*	User goal	<turn-on-television> <today's-weather>
Actions, *A*	System response	<show-today's-weather> <ask-user-to-repeat>
Observations, *Z*	Speech recognition hypotheses	<"what is today's weather">, confidence: 0.8
Belief, *b*	Probability mass function over states	{ 'turn-on-television': 0.9, 'today's-weather': 0.1 }

Function	Description	Example
Transition function, *T(g' | g, a)*	Effect of system action on user goal	$T(g_i \mid g_j, a) = 1$, if i == j $T(g_i \mid g_j, a) = 0$, if i != j
Observation function, *Ω(z| g, a)*	Mapping of ASR hypotheses to user intents	*Ω*("yes" | <today's-weather> , <confirm-today's-weather>) = *0.9*
Reward function, *R(g, a)*	Rewards for state-action pairs	R(<today's-weather>, <show-today's-weather>) = +10 R(<today's-weather>, <turn-on-television>) = -100
Policy, *Π()*	Mapping of beliefs to system actions	*Π(*[P(<today's-weather> = 0.99, …]) = <show-today's-weather>

Belief update: $b_T(g') = \eta \cdot \Omega(z|g_{T-1}, a) \sum_g T(g'|g,a)b(g)$



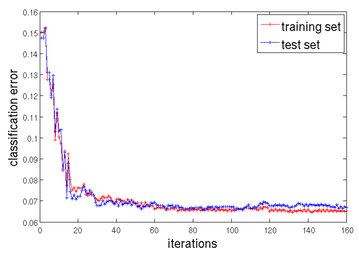


Belief state visualization (left) and user interface (right)

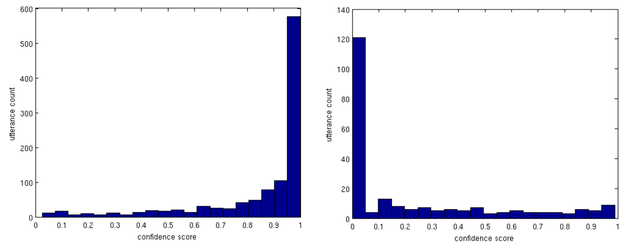
- Formulation as partially observable Markov decision process (POMDP): The agent (dialog manager) receives noisy observations (speech recognition hypotheses) of the hidden state (user's goal) and chooses actions (dialog responses)
- System functions: Information (weather, activities schedules, and menus) and communication (hands-free Skype calls)

Utterance Confidence Scoring

- Confidence scores encode the probability of hypothesis correctness
- Adaptive boosting (AdaBoost) used on annotated dataset of 2701 utterances for feature selection
- Acoustic, language-model, syntactic, semantic, and corpus-level features



Adaboost performance by iteration



Histograms of correct (left) and incorrect (right) hypotheses

- Observation model can then be expressed as:

$$\Omega(z_d, z_c|s,a) = P(z_d|s,a)\, P(z_c|z_d,s,a)$$

$$P(z_c|z_d,s,a) = \begin{cases} P(z_c|\text{correct}), z_d \mapsto s \\ P(z_c|\text{incorrect}), \text{otherwise} \end{cases}$$

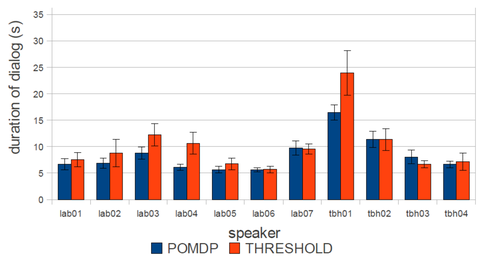
Experimental Results

- Comparison of threshold- and POMDP-based dialog systems using Ω(z| g, a) as simulated user model

Comparison of Dialog Strategies (50000 simulated dialogs)

	Dialog Completion			Number of Turns	
Dialog Manager	correct	incorrect	error rate	turns	std. error
argmax(State)	45920	4080	8.16%	1.07	0.02
threshold = 0.50	46589	3411	6.82%	2.60	0.07
threshold = 0.85	46976	3024	6.05%	5.44	0.14
threshold = 0.95	46179	3821	7.64%	10.39	0.17
POMDP-confidence	48022	1978	3.96%	1.23	0.00
POMDP+confidence	49421	579	1.16%	1.95	0.01

- Experiments with human participants:



Completion times for 20 dialog scenarios

Dialog Completion Rates for TBH speakers

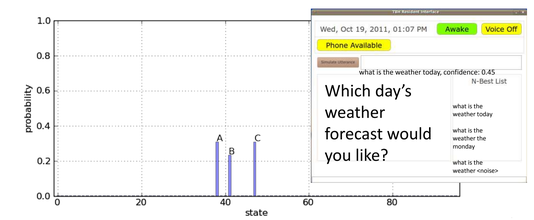
	completed dialogs (/20)	
	POMDP	THRESHOLD
tbh01	18	13
tbh02	17	16
tbh03	20	20
tbh04	19	18
tbh05	13	5
tbh06	18	10
tbh07	17	10

Conclusions

- End-to-end spoken dialog system designed to handle challenging speech input
- SDS-POMDP with confidence scoring outperforms threshold-based approaches for speakers with high error rates

Possible future directions:

- Larger/factored state space/system functions (*e.g.* robot mobility and manipulation)
- More nuanced clarification questions (using word-or sub-concept-level confidence scoring)



|Legend: A: ['weather', ['today']]; B: ['weather', ['monday']]; C: ['weather', ['sunday']]