Using Participant Role in Multiparty Meetings as Prior Knowledge for Nonparametric Topic Modeling

Songfang Huang and Steve Renals, CSTR, University of Edinburgh, {s.f.huang, s.renals}@ed.ac.uk

1. Background

AMI and AMIDA Projects:

- Instrumented and interactive environments
- Recognition, analysis, and modeling of communication
- Provide algorithms, models, and prototype systems that support interaction in meetings and access to meeting-related information
- Support meetings whether remote or co-located, real-time or asynchronous

AMI Meeting Corpus:

- 100 hours of multimodal meeting recordings
- Comprehensive annotations at a number of levels
- About 70% was elicited using a design scenario

Motivation:

- A scenario meeting in which the four participants (with different roles) decide to develop a new type of television remote control:

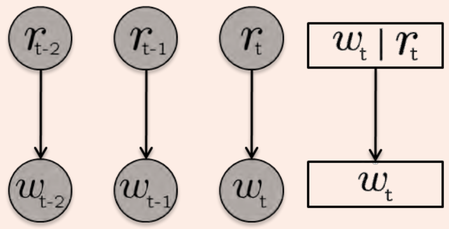
- Intuition: different words specific to each role?
  - •Project Manager (PM): *meeting*, *project*, …
  - •Marketing Expert (ME): *market*, *product*, …
  - •User Interface Designer (UI): *LCD*, *voice*, *speech*, …
  - •Industrial Designer (ID): *battery*, *chip*, *infrared*, …
- We are interested in:
  - •The word distribution over the role P(w|r)
  - •Will the role be helpful for topic modeling?
  - •Will the role be helpful for automatic speech recognition?

2. Modeling Approaches

Deterministic Approach (role-FLM):

Factored Language Model

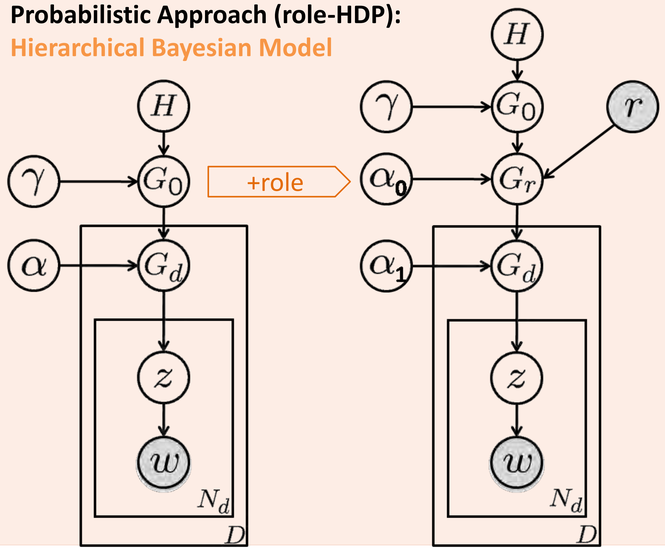
$$w_t \equiv f_t^{1:K} = \{f_t^1, f_t^2, \ldots, f_t^K\}$$

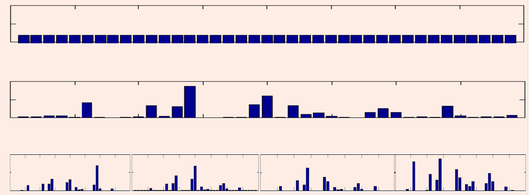


$$P(w|r) = \text{count}(w, r)/\text{count}(w)$$

Probabilistic Approach (role-HDP):

Hierarchical Bayesian Model



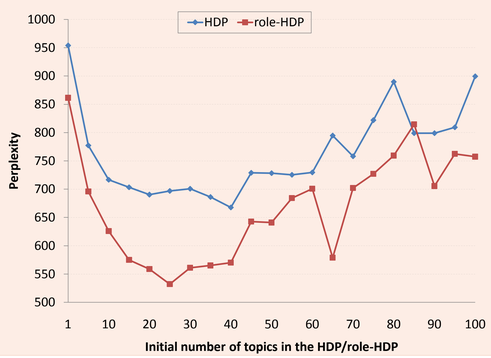


PM		ME		UI		ID	
(29) 0.28	(14) 0.13	(20) 0.27	(14) 0.16	(14) 0.25	(19) 0.15	(14) 0.17	(6) 0.16
DESIGN	BUTTON	REMOTE	BUTTON	BUTTON	RECOGNITION	BUTTON	CHIP
MEETING	BUTTONS	CONTROL	BUTTONS	CHANNEL	SPEECH	CHANNEL	SIMPLE
PROJECT	CHANNEL	LOOK	CHANNEL	SCREEN	VOICE	SCREEN	ADVANCED
MINUTES	SCREEN	MARKET	SCREEN	SCROLL	L_C_D_	SCROLL	REGULAR
USER	SCROLL	CONTROLS	SCROLL	VOLUME	SCREEN	VOLUME	REMOTE
INTERFACE	VOLUME	MEAN	VOLUME	MENU	MICROPHONE	MENU	L_C_D_
PRESENT	MENU	EASY	MENU	L_C_D_	CONTROLLER	L_C_D_	BATTERY
PRODUCT	L_C_D_	PRODUCT	L_C_D_	REMOTE	REMOTE	REMOTE	SPEAKER
DESIGNER	REMOTE	DESIGN	REMOTE	WHEEL	PROBLEM	WHEEL	EXPENSIVE
LOOK	WHEEL	BUTTONS	WHEEL	MEAN	TECHNOLOGY	MEAN	INFRARED
START	MEAN	FANCY	MEAN	CHANNELS	COFFEE	CHANNELS	STATION
SURE	CHANNELS	IMPORTANT	CHANNELS	CONTROL	SPEAKER	CONTROL	SAMPLE
BIT	CONTROL	PERCENT	CONTROL	PRESS	COST	PRESS	BATTERIES
GUESS	PRESS	FIND	PRESS	PRESS	CONTROL	PRESS	COST
THANK	KIND	LOT	KIND	KIND	SYSTEM	KIND	CONTROL

$$P(w|r) \approx \sum_{k=1}^{K} \phi_{kw} \cdot \theta_{dk}$$

3. Perplexity Results

- Data: fold-2–4 in the AMI Meeting Corpus for the training; fold1 for the testing.
- A vocabulary of 7,910 words.
- MCMC sampling for inference, auxiliary variable sampling for concentration parameters.
- 3,000 iterations to burn-in, 10 samples from posteriors, sample step 5.



4. Automatic Speech Recognition Results

- Data: 118 scenario meetings in the AMI Meeting Corpus for the training; the other 20 (~11 hours) for the testing.
- Baseline: HMM-based ASR – AMIASR (*Hain, et al.,2007*).
- A vocabulary with 56,168 words.
- Two baseline LMs trained on: 1) Fisher; 2) AMI+Fisher+HUB4
- P(w|r) for adaptation: $P_{\text{adapt}}(w|h) = P_{\text{back}}(w|h) \cdot \left(\frac{P(w|r)}{P_{\text{back}}(w)}\right)^{\mu} / z(h)$
- Baseline and adapted LMs were used for rescoring 500-best lists obtained from the lattices.

LMs	SUB	DEL	INS	WER
Fisher	22.7	11.4	5.8	**39.9**
role-FLM-adapted	22.5	11.1	5.9	**39.5**
HDP-adapted	22.2	11.3	5.6	**39.1**
role-HDP-adapted	22.3	11.3	5.6	**39.2**
AMI+Fisher+HUB4	21.6	11.1	5.4	**38.2**
role-FLM-adapted	21.4	10.9	5.6	**37.9**
HDP-adapted	21.2	11.1	5.3	**37.6**
role-HDP-adapted	21.2	11.1	5.3	**37.5**

5. Discussion

- **Probabilistic.** Each document is regarded as a multinomial distribution over roles, and each role a multinomial distribution over topics.
- **Observed vs. Latent.** Observed: to exploit the role as prior knowledge for topic modeling; Latent: to use other information to infer the roles for each document.
- **Application.** 1) The reduction in perplexity does not necessarily translate into the reduction in word error rate. 2) We observed no significant difference between the HDP and the role-HDP for adapting LMs in ASR. 3) Explicitly conditioning on the role for LMs, or tightly combining topic models and n-gram LMs.

6. Conclusion

- We observed that the role-HDP provides significant reductions in perplexity comparing to the HDP, and in word error rate comparing to the role-FLM.

Workshop on Prior Knowledge for Text and Language Process at ICML/UAI/CILT 2008, Helsinki, Finland
This work is jointly supported by the Wolfson Microelectronics Scholarship and the European IST Programme Project FP6-033812 (AMIDA).