System Combination with Log-linear Models

Jingzhou Yang, Chao Zhang, Anton Ragni, Mark Gales and Phil Woodland

Department of Engineering, University of Cambridge

{*jy308, cz277, ar527, mjfg, pcw*}*@eng.cam.ac.uk*

1 Introduction

- The commonly used system combination approaches:
  - Hypothesis combination:
    - * Recogniser output voting error reduction (ROVER).
    - * Confusion network combination (CNC).

      Need Multiple passes of decoding.
  - Log-likelihood combination:
    - * Joint decoding.

      Need single pass of decoding.

2 Joint Decoding

- The systems to be combined have the same HMM topology.
- Log-likelihoods are combined at frame level.
  - Combine hybrid and tandem systems:

$$\mathcal{L}(\boldsymbol{o}_t|s_i) \propto \eta_{\mathrm{H}} \underbrace{\log p_{\mathrm{H}}(\boldsymbol{o}_t|s_i)}_{\text{hybrid log-likelihood}} + \eta_{\mathrm{T}} \underbrace{\log p_{\mathrm{T}}(\boldsymbol{o}_t|s_i)}_{\text{tandem log-likelihood}}$$

  $\eta_{\mathrm{H}} = 1.0$ and $\eta_{\mathrm{T}} = 0.25$: combination weights used for hybrid and tandem systems.
- The combined log-likelihoods are un-normalised.
- Decoding:
  - Viterbi decoding is used.
  - Generated lattices are suitable for lattice rescoring.
- Related to log-linear models (LLMs).
  - Cache arc likelihoods in lattice for efficient rescoring.

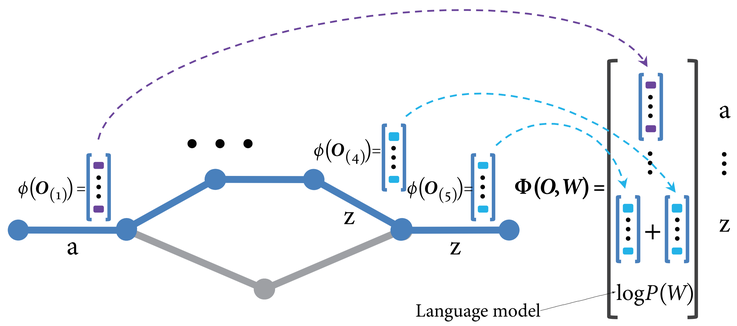
3 Structured Log-linear Models

- Systems can be combined at segment level.
- Relax the frame level Markov assumption to segment level.
- Capture long-span dependencies within the segment.
- The conditional distribution of the word sequence W given the observations $\boldsymbol{O}$ is modelled:

$$P(W|\boldsymbol{O}, \boldsymbol{\eta}) \propto \exp\left(\boldsymbol{\eta}^{\mathrm{T}} \Phi(\boldsymbol{O}, W)\right)$$

$\boldsymbol{\eta}$: model parameters (weights); $\Phi(\boldsymbol{O}, W)$: feature vector.

- The form of the feature vector $\Phi(\boldsymbol{O}, W)$:

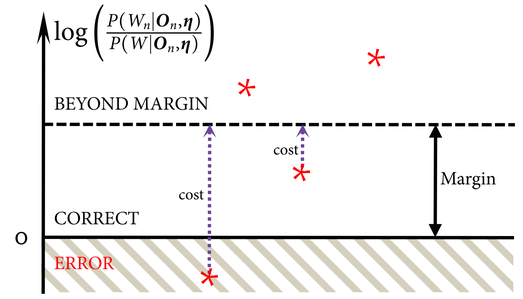


- Features for segment $\boldsymbol{O}_{(i)}$ with (tri-phone) hypothesis w_i:

$$\phi(\boldsymbol{O}_{(i)}) = \begin{bmatrix} \log p_{\mathrm{H}}(\boldsymbol{O}_{(i)}|w_i) \\ \log p_{\mathrm{T}}(\boldsymbol{O}_{(i)}|w_i) \end{bmatrix} \begin{matrix} \text{hybrid log-likelihood} \\ \text{tandem log-likelihood} \end{matrix}$$

4 Training and Decoding

- Large margin (LM):



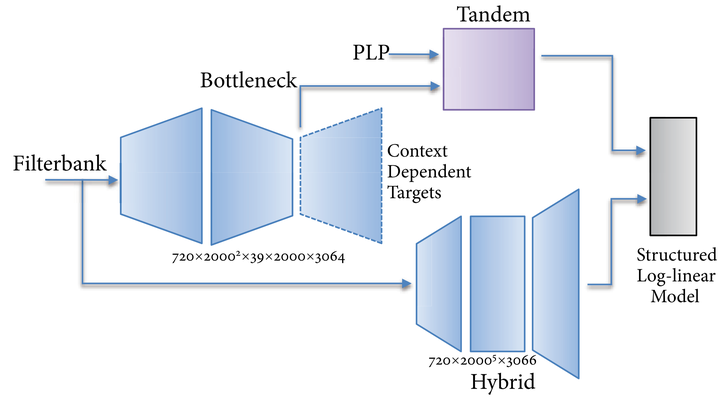
  - Margin: log-posterior ratio.
  - Introduce loss — minimise $\mathcal{F}_{\mathrm{LM}}(\boldsymbol{\eta})$:

$$\mathcal{F}_{\mathrm{LM}}(\boldsymbol{\eta}) = \sum_{n=1}^{N} \left[\max_{W \neq W_n} \left\{ \mathcal{L}(W, W_n) - \log\left(\frac{P(W_n|\boldsymbol{O}_n, \boldsymbol{\eta})}{P(W|\boldsymbol{O}_n, \boldsymbol{\eta})}\right)\right\}\right]_+$$

- Decoding:
  - Lattice rescoring, i.e. Viterbi algorithm applied to lattices.

5 Experiments

- The system framework:



  - Use 13d PLP and 40d log-Mel filter bank coefficients, and their first and higher order delta coefficients.
  - Use the standard bigram configuration.
- Results on the AURORA 4 dataset:

System	Criterion	Test Set WER(%)				Avg.
		Set A	Set B	Set C	Set D	
Tandem	MPE	4.78	7.63	8.93	19.14	12.45
Hybrid		3.75	6.70	7.68	17.62	11.24
CNC	–	3.87	6.76	7.45	17.17	11.06
Joint	Empirical	3.79	6.47	7.86	17.34	11.04
LLM	Empirical	3.74	6.57	7.88	17.12	10.98
	Large Margin	3.64	6.56	7.04	16.83	10.79

6 Conclusions

- Structured log-linear models
  - Relax the Markov assumption to segment level.
  - Use features from multiple systems.
    - * The combination weights are trained.

7 Acknowledgement

This work was supported in part by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Defense U.S. Army Research Laboratory (DoD / ARL) contract number W911NF-12-C-0012. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoD/ARL, or the U.S. Government.