

# A\* based Joint Segmentation and Classification of Dialog Acts in Multiparty Meetings

Matthias Zimmermann<sup>1</sup>, Yang Liu<sup>1,2</sup>, Elizabeth Shriberg<sup>1,3</sup>, Andreas Stolcke<sup>1,3</sup>

<sup>1</sup>International Computer Science Institute (ICSI), Berkeley, USA, <sup>2</sup>University of Texas, Dallas, USA, <sup>3</sup>SRI International, Menlo Park, USA  
{zimmerma, yangl, ees, stolcke}@icsi.berkeley.edu

## Abstract

We investigate the A\* algorithm for joint segmentation and classification of dialog acts (DAs) of the ICSI Meeting Corpus motivated by [1]. For the heuristic search a probabilistic framework is used that is based on DA-specific N-gram language models.

Furthermore, new metrics for performance evaluation are described and the influence of different metrics for performance evaluation is demonstrated.

The proposed method is evaluated on both traditional and new metrics, and compared with our previous work on the same task.

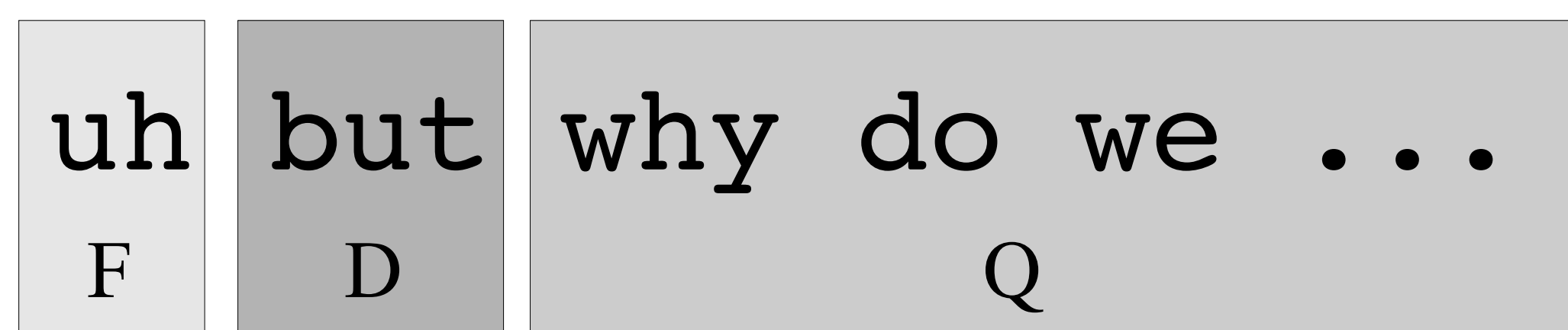
## Task

The task at hand is to split a stream of words into non-overlapping segments of text that are assigned 5 mutually exclusive DA labels B, D, F, Q, and S:

Input: Word Stream

uh but why do we ...

Output: Dialog Act Segments and Classes



B=backchannel, D=disruption, F=floor grabber, Q=question, and S=statement

We thank Barbara Peskin for her valuable comments. This work was partly supported by the European Union 6th FWP IST Integrated Project AMI (Augmented Multi-party Interaction, FP6-506811, publication), by DARPA Contract NBCHD030010 through the SRI CALO project (approved for public release, distribution unlimited), NSF Awards IIS-0121396 and IRI-9619921, and the Swiss National Science Foundation through the research network IM2.

## Metrics

For the assessment of the joint performance of the segmentation and classification of DAs, a word-based “**Strict**” and a DA-based “**DER**” metric are used in the experiments.

Ref: S | Q . Q . Q . Q | S . S . S | B | S . S  
 Sys: S | Q | S | Q . Q | D . D . D | S . S | S  
 Word based errors: c e e e e e e e e e e  
 DA based errors: c | e | e | e | e

Metric	Errors	Reference	Error Rate
<b>Strict</b>	10 match errors	11 words	91%
<b>DER</b>	4 match errors	5 DAs	80%

## Methodology

Assume  $n$  input words  $W = (w_1 \dots w_n)$ , we try to find a segmentation  $S = (s_1 \dots s_m)$  with corresponding DA labels  $D = (d_1 \dots d_m)$ . The  $s_i$  define the number of consecutive words in the  $i^{\text{th}}$  DA  $W_i = (w_k \dots w_l)$ , where  $s_i = w_l - w_k + 1$

Joint segmentation and classification of DAs can now be formulated as:

$$(\hat{S}, \hat{D}) = \arg \max_{S, D} P(S, D | W)$$

Bayes' rule and some independence assumptions lead to the product given below:

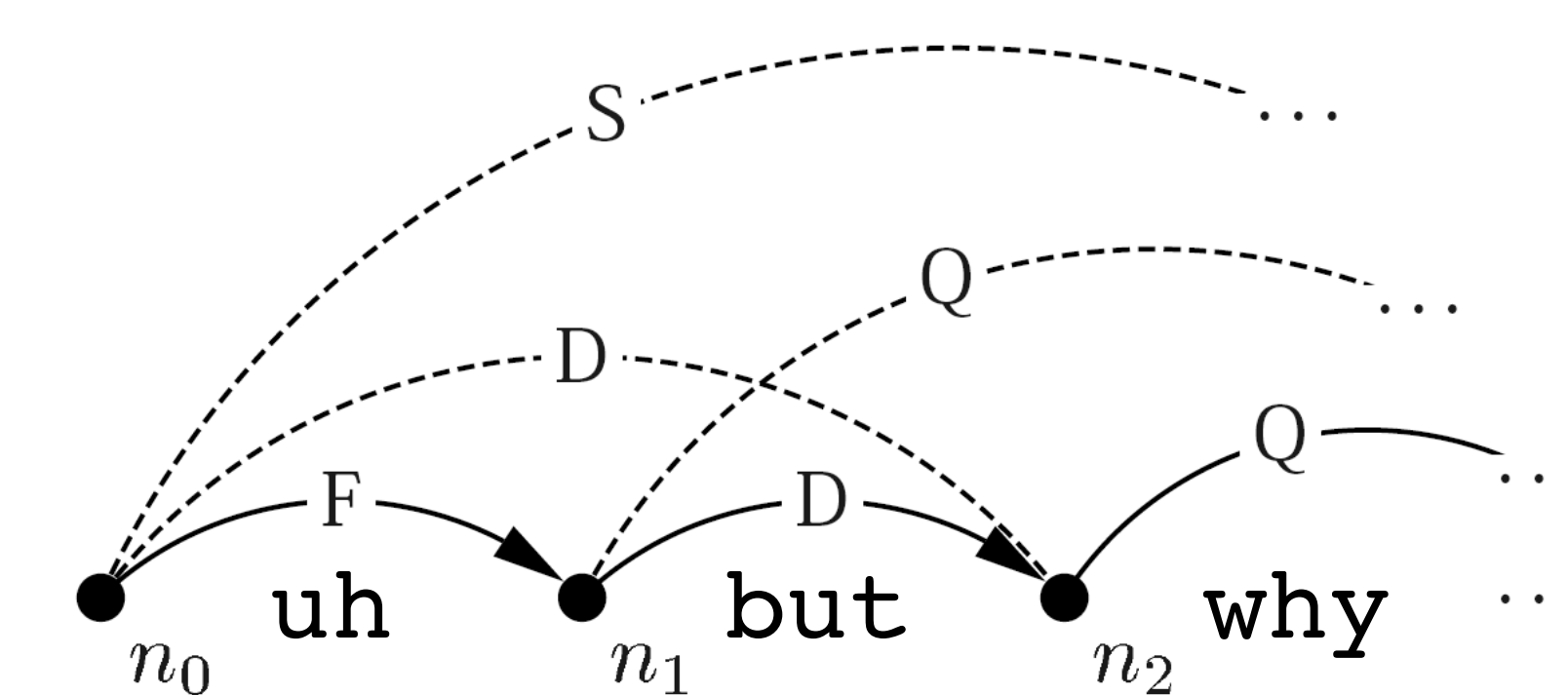
$$\prod_{i=1}^m P(W_i | s_i, d_i) P(s_i | d_i) P(d_i | d_{i-1})$$

that can be mapped to the following cost function:

$$\mathcal{C}_i = -\lambda_1 \log P(W_i | s_i, d_i) - \lambda_2 \log P(s_i | d_i) - \lambda_3 \log P(d_i | d_{i-1})$$

DA specific trigram language models are used for  $P(W_i | s_i, d_i)$ , the  $P(s_i | d_i)$  are computed from DA specific lengths, and a DA bigram grammar represents  $P(d_i | d_{i-1})$ .

The cost function is then used in the A\* graph search to find the most promising segmentation and classification



## Experiments

The experimental setup of [2] was used. To better compare results the original system of [2] has also been reduced to a system [2] np without prosody. Results are reported for both reference (Ref) and speech to text (STT) conditions:

Cond.	System	Strict	DER
Ref	[2]	64.4%	54.4%
	[2] np	72.4%	64.1%
	A*	73.1%	62.3%
STT	[2]	75.4%	64.3%
	[2] np	82.9%	73.2%
	A*	83.9%	71.4%

## Conclusion

Joint segmentation and classification of DA based on A\* seems to be promising. Next steps will include the use of prosody and processing of word lattices.

[1] V. Warnke et al. “Integrated dialog act segmentation and classification using prosodic features and language models”, EUROSPEECH, 1997

[2] J. Ang et al. “Automatic dialog act segmentation and classification in multiparty meetings”, ICASSP, 2005