Toward Joint Segmentation and Classification of Dialog Acts in Multiparty Meetings

MLMI'05, July 11-13th

Matthias Zimmermann, Yang Liu, Elizabeth Shriberg, and Andreas Stolcke

{zimmerma,yang,ees,stolcke}@icsi.berkeley.edu

International Computer Science Institute, Berkeley, USA

Organization

1. Problem Statement
2. Hidden-Event LM and Tagger
3. Performance Metrics
4. Experiments and Results
5. Conclusions and Outlook

Problem Statement

Segmentation of a multiparty meeting into its Dialog Acts (DAs)

Problem Statement

Segmentation of a multiparty meeting into its Dialog Acts (DAs)

Input: well u- that's pretty good i think yeah thanks

Problem Statement

Segmentation of a multiparty meeting into its Dialog Acts (DAs)

Input:	well u- that's pretty good i think yeah thanks
Transcript:	$[_D$ well$]$ $[_S$ u- that's pretty good i think$]$ $[_S$ yeah$]$ $[_S$ thanks$]$

Problem Statement

Segmentation of a multiparty meeting into its Dialog Acts (DAs)

Input:	well u- that's pretty good i think yeah thanks
Transcript:	$[_D$ well$]$ $[_S$ u- that's pretty good i think$]$ $[_S$ yeah$]$ $[_S$ thanks$]$
System:	$[_S$ well u- that's pretty good$]$ $[_D$ i think$]$ $[_B$ yeah$]$ $[_S$ thanks$]$

Problem Statement

Previous work

Problem Statement

Previous work

- typically either segmentation or classification of DAs

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

1997, Warnke et al.

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

1997, Warnke et al.

- Verbmobil

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

1997, Warnke et al.

- Verbmobil
- integrated segmentation and classification

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

1997, Warnke et al.

- Verbmobil
- integrated segmentation and classification

2005, Ang et al. (ICASSP'05)

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

1997, Warnke et al.

- Verbmobil
- integrated segmentation and classification

2005, Ang et al. (ICASSP'05)

- ICSI meeting corpus

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

1997, Warnke et al.

- Verbmobil
- integrated segmentation and classification

2005, Ang et al. (ICASSP'05)

- ICSI meeting corpus
- sequential approach

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

1997, Warnke et al.

- Verbmobil
- integrated segmentation and classification

2005, Ang et al. (ICASSP'05)

- ICSI meeting corpus
- sequential approach
- segmentation into DAs: hidden event LM, and decision trees

Problem Statement

Previous work

- typically either segmentation or classification of DAs
- fully automatic systems require solutions to both

1997, Warnke et al.

- Verbmobil
- integrated segmentation and classification

2005, Ang et al. (ICASSP'05)

- ICSI meeting corpus
- sequential approach
- segmentation into DAs: hidden event LM, and decision trees
- classification of DAs: maximum entropy, and decision trees

Hidden Event Language Model

Hidden Event LM (HE-LM)

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events
- hidden events typically correspond to sentence boundaries

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events
- hidden events typically correspond to sentence boundaries
- in our case events act as tagged DA boundaries

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events
- hidden events typically correspond to sentence boundaries
- in our case events act as tagged DA boundaries

Training: well <d> u- that's pretty good i think <s> yeah <b> thanks <s>

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events
- hidden events typically correspond to sentence boundaries
- in our case events act as tagged DA boundaries

Training: well <d> u- that's pretty good i think <s> yeah <b> thanks <s>

Testing: well u- that's pretty good i think yeah thanks

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events
- hidden events typically correspond to sentence boundaries
- in our case events act as tagged DA boundaries

Training: well <d> u- that's pretty good i think <s> yeah <b> thanks <s>

Testing: well u- that's pretty good i think yeah thanks

$$p(\text{<b>}|\text{pretty}), p(\text{<d>}|\text{pretty}), \ldots p(\text{<>}|\text{pretty})$$

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events
- hidden events typically correspond to sentence boundaries
- in our case events act as tagged DA boundaries

Training: well <d> u- that's pretty good i think <s> yeah <b> thanks <s>

Testing: well u- that's pretty good i think yeah thanks

$$p(\text{<b>}|\text{pretty}), p(\text{<d>}|\text{pretty}), \ldots p(\text{<>}|\text{pretty})$$

- after each word the event with the highest posterior is inserted

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events
- hidden events typically correspond to sentence boundaries
- in our case events act as tagged DA boundaries

Training: well <d> u- that's pretty good i think <s> yeah <b> thanks <s>

Testing: well u- that's pretty good i think yeah thanks

$$p(\text{<b>}|\text{pretty}), p(\text{<d>}|\text{pretty}), \ldots p(\text{<>}|\text{pretty})$$

- after each word the event with the highest posterior is inserted
- all words to the left of a DA boundary event inherit its label

Hidden Event Language Model

Hidden Event LM (HE-LM)

- N-gram modeling for a stream of words including hidden events
- hidden events typically correspond to sentence boundaries
- in our case events act as tagged DA boundaries

Training: well <d> u- that's pretty good i think <s> yeah <b> thanks <s>

Testing: well u- that's pretty good i think yeah thanks

$$p(\text{<b>}|\text{pretty}), p(\text{<d>}|\text{pretty}), \ldots p(\text{<>}|\text{pretty})$$

- after each word the event with the highest posterior is inserted
- all words to the left of a DA boundary event inherit its label

Result: well$_s$ u-$_s$ that's$_s$ pretty$_s$ good$_{s|}$ i$_d$ think$_{d|}$ yeah$_{b|}$ thanks$_{s|}$

Tagger

Tagger

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T
- system tries to find V_T sequence with the highest posterior given:

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T
- system tries to find V_T sequence with the highest posterior given:
  - sequence of words in V

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T
- system tries to find V_T sequence with the highest posterior given:
  - sequence of words in V
  - mapping probabilities from words in V to words in V_T

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T
- system tries to find V_T sequence with the highest posterior given:
  - sequence of words in V
  - mapping probabilities from words in V to words in V_T
  - N-gram LM for sequence of words in V_T

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T
- system tries to find V_T sequence with the highest posterior given:
  - sequence of words in V
  - mapping probabilities from words in V to words in V_T
  - N-gram LM for sequence of words in V_T

Input in V well u- that's pretty good i think yeah thanks

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T
- system tries to find V_T sequence with the highest posterior given:
  - sequence of words in V
  - mapping probabilities from words in V to words in V_T
  - N-gram LM for sequence of words in V_T

Input in V	well u- that's pretty good i think yeah thanks
Mapping	$p(\text{yeah}\|\text{yeah}_{b+})$, $p(\text{yeah}\|\text{yeah}_{b})$, $p(\text{yeah}\|\text{yeah}_{d+})$, ...

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T
- system tries to find V_T sequence with the highest posterior given:
  - sequence of words in V
  - mapping probabilities from words in V to words in V_T
  - N-gram LM for sequence of words in V_T

Input in V	well u- that's pretty good i think yeah thanks
Mapping	$p($yeah$\vert$yeah$_{b+})$, $p($yeah$\vert$yeah$_{b})$, $p($yeah$\vert$yeah$_{d+})$, ...
LM in V_T	$p($yeah$_{s+}\vert$ i$_{d+}$, think$_{d})$

Tagger

Tagger

- translation of a stream of words from vocabulary V into words from a (tagged) vocabulary V_T
- system tries to find V_T sequence with the highest posterior given:
  - sequence of words in V
  - mapping probabilities from words in V to words in V_T
  - N-gram LM for sequence of words in V_T

Input in V	well u- that's pretty good i think yeah thanks
Mapping	$p($yeah$\|$yeah$_{b+})$, $p($yeah$\|$yeah$_{b})$, $p($yeah$\|$yeah$_{d+})$, ...
LM in V_T	$p($yeah$_{s+}\|$ i$_{d+}$, think$_{d})$

Result in V_T	well$_{s+}$ u-$_{s}$ that's$_{s}$ pretty$_{s}$ good$_{s}$ i$_{d+}$ think$_{d}$ yeah$_{s+}$ thanks$_{s}$

Performance Metrics

Previous Metrics

Performance Metrics

Previous Metrics

- boundary based: NIST-SU metrics

Performance Metrics

Previous Metrics

- boundary based: NIST-SU metrics
- word based: Lenient, and Strict metrics (ICASSP'05)

Performance Metrics

Previous Metrics

- boundary based: NIST-SU metrics
- word based: Lenient, and Strict metrics (ICASSP'05)

Proposed Metrics: DA Based

Performance Metrics

Previous Metrics

- boundary based: NIST-SU metrics
- word based: Lenient, and Strict metrics (ICASSP'05)

Proposed Metrics: DA Based

- simple to interprete, directly related to DAs

Performance Metrics

Previous Metrics

- boundary based: NIST-SU metrics
- word based: Lenient, and Strict metrics (ICASSP'05)

Proposed Metrics: DA Based

- simple to interprete, directly related to DAs
- counting units are the DAs as in transcripts

Performance Metrics

Previous Metrics

- boundary based: NIST-SU metrics
- word based: Lenient, and Strict metrics (ICASSP'05)

Proposed Metrics: DA Based

- simple to interprete, directly related to DAs
- counting units are the DAs as in transcripts
- percentage of wrongly segmented DAs:
  Dialog act Segmentation Error Rate (DSER)

Performance Metrics

Previous Metrics

- boundary based: NIST-SU metrics
- word based: Lenient, and Strict metrics (ICASSP'05)

Proposed Metrics: DA Based

- simple to interprete, directly related to DAs
- counting units are the DAs as in transcripts
- percentage of wrongly segmented DAs:
  Dialog act Segmentation Error Rate (DSER)
- percentage of wrongly segmented or classified DAs:
  Dialog act Error Rate (DER)

Segmentation Metrics

Metrics for Segmentation Errors

Segmentation Metrics

Metrics for Segmentation Errors

- NIST-SU, boundary based

$$NIST - SU = \frac{Misses + FA}{Boundaries} \times 100\%$$

Segmentation Metrics

Metrics for Segmentation Errors

- NIST-SU, boundary based

$$NIST - SU = \frac{Misses + FA}{Boundaries} \times 100\%$$

- Da Segmentation Error Rate (DSER), DA based

$$DER = \frac{missegmented\ \ DAs}{DAs} \times 100\%$$

Segmentation Metrics

Examples

Reference	`S|Q.Q.Q.Q|S.S.S|B|S.S|`
System	`S|Q|S|Q.Q|D.D.D|S.S.S|`

Segmentation Metrics

Examples

```
Reference    S|Q.Q.Q.Q|S.S.S|B|S.S|
System       S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU       C E E   C     C E   C
```

Segmentation Metrics

Examples

```
Reference    S|Q.Q.Q.Q|S.S.S|B|S.S|
System       S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU       C E E   C     C E   C
DSER         C|   E   |  C  |E| E |
```

Segmentation Metrics

Examples

```
Reference   S|Q.Q.Q.Q|S.S.S|B|S.S|
System      S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU      C E E   C     C E   C
DSER        C|   E   |  C  |E| E |
```

Metric	Errors	Reference Units	Error Rate
NIST-SU	2 FA, 1 miss	5 boundaries	60%

Segmentation Metrics

Examples

```
Reference   S|Q.Q.Q.Q|S.S.S|B|S.S|
System      S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU      C E E   C     C E   C
DSER        C|   E   |  C  |E| E |
```

Metric	Errors	Reference Units	Error Rate
NIST-SU	2 FA, 1 miss	5 boundaries	60%
DSER	3 match errors	5 DAs	60%

Segmentation and Classification Metrics

Metrics for Segmentation and Classification errors

Segmentation and Classification Metrics

Metrics for Segmentation and Classification errors

- NIST-SU, boundary based

$$NIST-SU = \frac{Substitutes + Misses + FA}{Boundaries} \times 100\%$$

Segmentation and Classification Metrics

Metrics for Segmentation and Classification errors

- NIST-SU, boundary based

$$NIST-SU = \frac{Substitutes + Misses + FA}{Boundaries} \times 100\%$$

- Lenient, word based (does not consider segmentation)

$$Lenient = \frac{mistagged\ Words}{Words} \times 100\%$$

Segmentation and Classification Metrics

- Strict, word based

$$Strict = \frac{mistagged\ \ or\ \ missegmented\ \ Words}{Words} \times 100\%$$

Segmentation and Classification Metrics

- Strict, word based

$$Strict = \frac{mistagged\ \ or\ \ missegmented\ \ Words}{Words} \times 100\%$$

- DA Error Rate (DER), DA based

$$DER = \frac{DAs\ \ containing\ \ errors}{DAs} \times 100\%$$

Segmentation and Classification Metrics

Examples

Reference	`S|Q.Q.Q.Q|S.S.S|B|S.S|`
System	`S|Q|S|Q.Q|D.D.D|S.S.S|`

Segmentation and Classification Metrics

Examples

```
Reference    S|Q.Q.Q.Q|S.S.S|B|S.S|
System       S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU       C E E   C     E E   C
```

Segmentation and Classification Metrics

Examples

```
Reference    S|Q.Q.Q.Q|S.S.S|B|S.S|
System       S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU       C E E   C     E E   C
Lenient      C C E C C E E E E C C
```

Segmentation and Classification Metrics

Examples

```
Reference   S|Q.Q.Q.Q|S.S.S|B|S.S|
System      S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU      C E E   C     E E   C
Lenient     C C E C C E E E E C C
Strict      C E E E E E E E E E E
```

Segmentation and Classification Metrics

Examples

```
Reference   S|Q.Q.Q.Q|S.S.S|B|S.S|
System      S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU      C E E   C     E E   C
Lenient     C C E C C E E E E C C
Strict      C E E E E E E E E E E
DER         C|   E   |  E  |E| E |
```

Segmentation and Classification Metrics

Examples

```
Reference   S|Q.Q.Q.Q|S.S.S|B|S.S|
System      S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU      C E E   C     E E   C
Lenient     C C E C C E E E E C C
Strict      C E E E E E E E E E E
DER         C|   E   |  E  |E| E |
```

Metric	Errors	Reference Units	Error Rate
NIST-SU	1 sub., 2 FA, 1 miss	5 boundaries	80%

Segmentation and Classification Metrics

Examples

```
Reference   S|Q.Q.Q.Q|S.S.S|B|S.S|
System      S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU      C E E   C     E E   C
Lenient     C C E C C E E E E C C
Strict      C E E E E E E E E E E
DER         C|   E   |  E  |E| E |
```

Metric	Errors	Reference Units	Error Rate
NIST-SU	1 sub., 2 FA, 1 miss	5 boundaries	80%
Lenient	5 match errors	11 words	45%

Segmentation and Classification Metrics

Examples

```
Reference   S|Q.Q.Q.Q|S.S.S|B|S.S|
System      S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU      C E E   C     E E   C
Lenient     C C E C C E E E E C C
Strict      C E E E E E E E E E E
DER         C|   E   |  E  |E| E |
```

Metric	Errors	Reference Units	Error Rate
NIST-SU	1 sub., 2 FA, 1 miss	5 boundaries	80%
Lenient	5 match errors	11 words	45%
Strict	10 match errors	11 words	91%

Segmentation and Classification Metrics

Examples

```
Reference   S|Q.Q.Q.Q|S.S.S|B|S.S|
System      S|Q|S|Q.Q|D.D.D|S.S.S|
NIST-SU      C E E   C     E E   C
Lenient     C C E C C E E E E C C
Strict      C E E E E E E E E E E
DER         C|   E   |  E  |E| E |
```

Metric	Errors	Reference Units	Error Rate
NIST-SU	1 sub., 2 FA, 1 miss	5 boundaries	80%
Lenient	5 match errors	11 words	45%
Strict	10 match errors	11 words	91%
DER	4 match errors	5 DAs	80%

Experimental Setup

ICSI meeting corpus with DA annotations (MRDA)

Experimental Setup

ICSI meeting corpus with DA annotations (MRDA)

- as in Ang et al. (ICASSP'05)

Experimental Setup

ICSI meeting corpus with DA annotations (MRDA)

- as in Ang et al. (ICASSP'05)
- 51 meetings for training, 11 for validation, and 11 for testing

Experimental Setup

ICSI meeting corpus with DA annotations (MRDA)

- as in Ang et al. (ICASSP'05)
- 51 meetings for training, 11 for validation, and 11 for testing
- 2 conditions: reference text, and STT$*$ output

$*$: average WER: 39%, 32% for native speaker

Experimental Setup

ICSI meeting corpus with DA annotations (MRDA)

- as in Ang et al. (ICASSP'05)
- 51 meetings for training, 11 for validation, and 11 for testing
- 2 conditions: reference text, and STT* output
- 5 DA types†

$*$: average WER: 39%, 32% for native speaker

$\dagger$: B=Backchannel, D=Disruption, F=Floor grabber, Q=Question, S=Statement

Segmentation Performance

Condition	System	NIST-SU	DSER
Ref	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
STT			

ICASSP'05* without prosody features

Segmentation Performance

Condition	System	NIST-SU	DSER
	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
Ref	HE-LM	46.3	55.3
STT			

ICASSP'05* without prosody features

Segmentation Performance

Condition	System	NIST-SU	DSER
Ref	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
	HE-LM	46.3	55.3
	Tagger	51.1	61.7
STT			

ICASSP'05* without prosody features

Segmentation Performance

Condition	System	NIST-SU	DSER
Ref	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
	HE-LM	46.3	55.3
	Tagger	51.1	61.7
STT	ICASSP'05	45.5	49.4
	ICASSP'05*	59.5	62.0

ICASSP'05* without prosody features

Segmentation Performance

Condition	System	NIST-SU	DSER
	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
Ref	HE-LM	46.3	55.3
	Tagger	51.1	61.7
	ICASSP'05	45.5	49.4
	ICASSP'05*	59.5	62.0
STT	HE-LM	59.6	62.4

ICASSP'05* without prosody features

Segmentation Performance

Condition	System	NIST-SU	DSER
Ref	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
	HE-LM	46.3	55.3
	Tagger	51.1	61.7
STT	ICASSP'05	45.5	49.4
	ICASSP'05*	59.5	62.0
	HE-LM	59.6	62.4
	Tagger	62.8	66.9

ICASSP'05* without prosody features

Segmentation and Classification Performance

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
Ref	ICASSP'05*	62.3	21.0	72.4	64.1
STT					

ICASSP'05* without prosody features

Segmentation and Classification Performance

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
Ref	HE-LM	62.2	23.3	74.3	66.5
STT					

ICASSP'05* without prosody features

Segmentation and Classification Performance

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
Ref	HE-LM	62.2	23.3	74.3	66.5
	Tagger	69.5	22.6	78.6	72.6
STT					

ICASSP'05* without prosody features

Segmentation and Classification Performance

Condition	System	NIST-SU	Lenient	Strict	DER
Ref	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
	HE-LM	62.2	23.3	74.3	66.5
	Tagger	69.5	22.6	78.6	72.6
STT	ICASSP'05	68.3	25.1	75.4	64.3
	ICASSP'05*	78.3	25.0	82.9	73.2

ICASSP'05* without prosody features

Segmentation and Classification Performance

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP’05	52.6	20.0	64.4	54.4
	ICASSP’05*	62.3	21.0	72.4	64.1
Ref	HE-LM	62.2	23.3	74.3	66.5
	Tagger	69.5	22.6	78.6	72.6
	ICASSP’05	68.3	25.1	75.4	64.3
	ICASSP’05*	78.3	25.0	82.9	73.2
STT	HE-LM	78.0	26.2	83.8	73.9

ICASSP’05* without prosody features

Segmentation and Classification Performance

Condition	System	NIST-SU	Lenient	Strict	DER
Ref	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
	HE-LM	62.2	23.3	74.3	66.5
	Tagger	69.5	22.6	78.6	72.6
STT	ICASSP'05	68.3	25.1	75.4	64.3
	ICASSP'05*	78.3	25.0	82.9	73.2
	HE-LM	78.0	26.2	83.8	73.9
	Tagger	81.3	22.4	85.4	77.3

ICASSP'05* without prosody features

Conclusions

Main Conclusions

Conclusions

Main Conclusions

- investigated HE-LM and Tagger based approaches

Conclusions

Main Conclusions

- investigated HE-LM and Tagger based approaches
- established baseline for joint segmentation and classification

Conclusions

Main Conclusions

- investigated HE-LM and Tagger based approaches
- established baseline for joint segmentation and classification
- proposed and motivated DA based DSER and DER metrics

Outlook

Future Directions

Outlook

Future Directions

- A* algorithm to take into account complete DA hypotheses

Outlook

Future Directions

- A* algorithm to take into account complete DA hypotheses
- integrate both word and prosody based information

Outlook

Future Directions

- A* algorithm to take into account complete DA hypotheses
- integrate both word and prosody based information
- use of word lattices produced by STT

Thank You

more 1

Condition	System	NIST-SU	DSER
	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
Ref			
STT			

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 1

Condition	System	NIST-SU	DSER
	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
Ref	HE-LM	46.3	55.3
STT			

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 1

Condition	System	NIST-SU	DSER
	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
Ref	HE-LM	46.3	55.3
	A*	63.3	47.1
STT			

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 1

Condition	System	NIST-SU	DSER
	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
Ref	HE-LM	46.3	55.3
	A*	63.3	47.1
	A$^*_{DA}$	51.0	48.9
STT			

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 1

Condition	System	NIST-SU	DSER
Ref	ICASSP’05	34.5	40.8
	ICASSP’05*	46.0	53.0
	HE-LM	46.3	55.3
	A*	63.3	47.1
	A$^*_{DA}$	51.0	48.9
STT	ICASSP’05	45.5	49.4
	ICASSP’05*	59.5	62.0

ICASSP’05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 1

Condition	System	NIST-SU	DSER
Ref	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
	HE-LM	46.3	55.3
	A*	63.3	47.1
	A$^*_{DA}$	51.0	48.9
STT	ICASSP'05	45.5	49.4
	ICASSP'05*	59.5	62.0
	HE-LM	59.6	62.4

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 1

Condition	System	NIST-SU	DSER
Ref	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
	HE-LM	46.3	55.3
	A*	63.3	47.1
	A$^*_{DA}$	51.0	48.9
STT	ICASSP'05	45.5	49.4
	ICASSP'05*	59.5	62.0
	HE-LM	59.6	62.4
	A*	81.8	55.1

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 1

Condition	System	NIST-SU	DSER
Ref	ICASSP'05	34.5	40.8
	ICASSP'05*	46.0	53.0
	HE-LM	46.3	55.3
	A*	63.3	47.1
	A$^*_{DA}$	51.0	48.9
STT	ICASSP'05	45.5	49.4
	ICASSP'05*	59.5	62.0
	HE-LM	59.6	62.4
	A*	81.8	55.1
	A$^*_{DA}$	71.1	55.8

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 2

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
Ref					
STT					

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 2

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
Ref	HE-LM	62.2	23.3	74.3	66.5
STT					

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 2

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
Ref	HE-LM	62.2	23.3	74.3	66.5
	A*	90.1	28.7	76.5	63.2
STT					

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 2

Condition	System	NIST-SU	Lenient	Strict	DER
Ref	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
	HE-LM	62.2	23.3	74.3	66.5
	A*	90.1	28.7	76.5	63.2
	A$^*_{DA}$	72.0	24.7	73.2	62.3
STT					

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 2

Condition	System	NIST-SU	Lenient	Strict	DER
Ref	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
	HE-LM	62.2	23.3	74.3	66.5
	A*	90.1	28.7	76.5	63.2
	A$^*_{DA}$	72.0	24.7	73.2	62.3
STT	ICASSP'05	68.3	25.1	75.4	64.3
	ICASSP'05*	78.3	25.0	82.9	73.2

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 2

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
Ref	HE-LM	62.2	23.3	74.3	66.5
	A*	90.1	28.7	76.5	63.2
	A$^*_{DA}$	72.0	24.7	73.2	62.3
	ICASSP'05	68.3	25.1	75.4	64.3
	ICASSP'05*	78.3	25.0	82.9	73.2
STT	HE-LM	78.0	26.2	83.8	73.9

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 2

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
Ref	HE-LM	62.2	23.3	74.3	66.5
	A*	90.1	28.7	76.5	63.2
	A$^*_{DA}$	72.0	24.7	73.2	62.3
	ICASSP'05	68.3	25.1	75.4	64.3
	ICASSP'05*	78.3	25.0	82.9	73.2
STT	HE-LM	78.0	26.2	83.8	73.9
	A*	114.0	35.5	85.6	72.4

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM

more 2

Condition	System	NIST-SU	Lenient	Strict	DER
	ICASSP'05	52.6	20.0	64.4	54.4
	ICASSP'05*	62.3	21.0	72.4	64.1
Ref	HE-LM	62.2	23.3	74.3	66.5
	A*	90.1	28.7	76.5	63.2
	A$^*_{DA}$	72.0	24.7	73.2	62.3
	ICASSP'05	68.3	25.1	75.4	64.3
	ICASSP'05*	78.3	25.0	82.9	73.2
STT	HE-LM	78.0	26.2	83.8	73.9
	A*	114.0	35.5	85.6	72.4
	A$^*_{DA}$	98.4	31.5	83.9	71.4

ICASSP'05* without prosody features

A$^*_{DA}$ including DA bigram LM