On the Dynamics of Overlap in Multi-Party Conversation

Kornel Laskowski, Mattias Heldner & Jens Edlund

Carnegie Mellon University, Pittsburgh PA, USA
Voci Technologies, Inc., Pittsburgh PA, USA
Dept. of Linguistics, Stockholm University, Stockholm, Sweden
KTH Speech Music & Hearing, Stockholm, Sweden

11 September, 2012

What Is Overlap?

The occurrence of

more than one person speaking simultaneously.

What Is Overlap?

The occurrence of

more than one person speaking simultaneously.

An example ...

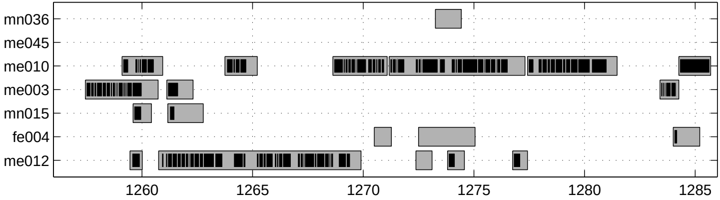
```
me003: Okay, so then I'll go back and look at the ones
       [on the l]ist [that -     ]
me010: [Okay.   ]    [And you can] ASK Kevin.
me012:              Y[eah.       ]
mn015:               [But -      ]
       (0.3)
me012: Yeah, the [one that] uh people seem to use =
me003:           [M[mm.   ]
mn015:             [But - ]
me012: = is uh Hugin or whatever?        [How exp-         ] =
me010:                            Hugin, [yeah that's free.]
me012: = I don't think it's - Is it free? Because I've seen it
       ADVERTISED in places so I - it [seems] [to -  ]
me010:                               U[h it ] [may be] free to
       academics. Like I - [I don't know.]
fe004:                     [((sniff))    ]
```

What Is Overlap?

The occurrence of

more than one person speaking simultaneously.

An example ...



a binary-valued speech/non-speech **chronogram**

The Occurrence of Overlap Confounds Systems

The Occurrence of Overlap Confounds Systems

1. The occurrence of overlap is acoustically difficult to detect.

The Occurrence of Overlap Confounds Systems

1. The occurrence of overlap is acoustically difficult to detect.
2. Simultaneous streams of speech are acoustically difficult to separate.

The Occurrence of Overlap Confounds Systems

1. The occurrence of overlap is acoustically difficult to detect.
2. Simultaneous streams of speech are acoustically difficult to separate.
3. Speech corrupted by other simultaneous speech is difficult to acoustically recognize.

The Occurrence of Overlap Confounds Systems

1. The occurrence of overlap is acoustically difficult to detect.
2. Simultaneous streams of speech are acoustically difficult to separate.
3. Speech corrupted by other simultaneous speech is difficult to acoustically recognize.
4. Speech deployed in overlap is grammatically distinct.

The Occurrence of Overlap Confounds Systems

1. The occurrence of overlap is acoustically difficult to detect.
2. Simultaneous streams of speech are acoustically difficult to separate.
3. Speech corrupted by other simultaneous speech is difficult to acoustically recognize.
4. Speech deployed in overlap is grammatically distinct.

It behooves us ...

- to seek to understand **when** it occurs
- to design methodologies for identifying it ...

The Occurrence of Overlap Confounds Systems

1. The occurrence of overlap is acoustically difficult to detect.
2. Simultaneous streams of speech are acoustically difficult to separate.
3. Speech corrupted by other simultaneous speech is difficult to acoustically recognize.
4. Speech deployed in overlap is grammatically distinct.

It behooves us ...

- to seek to understand **when** it occurs
- to design methodologies for identifying it ...
- ... at the earliest, lowest-level stage of processing

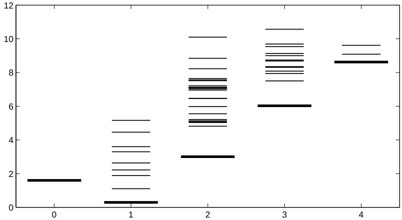
We Know **That** It Occurs ...

... and even **how frequently** various degree of overlap occur (Baron et al, 2001; Çetin et al, 2006)

We Know **That** It Occurs ...

... and even **how frequently** various degree of overlap occur (Baron et al, 2001; Çetin et al, 2006)

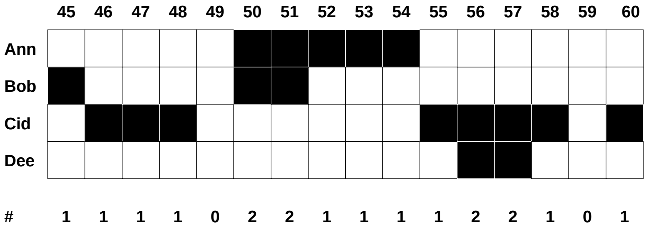
e.g. the negative log-probability of occurrence as a function of **degree-of-overlap** (Laskowski et al, 2010):



"degree-of-overlap" $\equiv$ number of simultaneously speaking participants

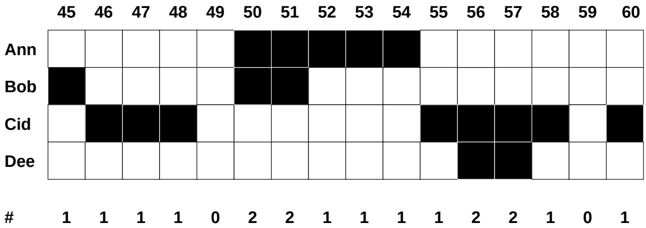
… But Not **When** It Occurs

If we were to take a chronogram and shuffle its time slices …



… But Not **When** It Occurs

If we were to take a chronogram and shuffle its time slices …



… we would get **the same prior probabilities** of occurrence.
Systems are currently at the mercy of these priors alone.

Plan for the Next 15 Minutes

FOCUS: the **sequence** of degree-of-overlap.

Plan for the Next 15 Minutes

FOCUS: the **sequence** of degree-of-overlap.

1. Ask a specific question about speech chronograms:

   "Do they look the same **right-to-left** as **left-to-right**?"

   1. Hypothesize that "H_1: They look different."
   2. Develop a stochastic modeling framework.
   3. Confidently reject H_0.

Plan for the Next 15 Minutes

FOCUS: the **sequence** of degree-of-overlap.

1. Ask a specific question about speech chronograms:

   "Do they look the same **right-to-left** as **left-to-right**?"

   1. Hypothesize that "H_1: They look different."
   2. Develop a stochastic modeling framework.
   3. Confidently reject H_0.

2. What **causes** this asymmetry?

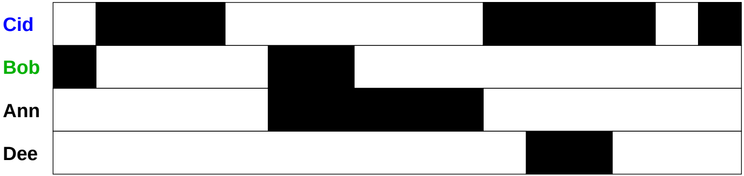
   1. Investigate what model learns.
   2. Investigate the effect of individual **dialog act (DA)** types ...
      ... by **ignoring** their contribution to overlap.
   3. Find that **only a handful** of DA types is responsible.

Modeling Degree-of-Overlap as a Stochastic Process

Modeling Degree-of-Overlap as a Stochastic Process

Modeling Degree-of-Overlap as a Stochastic Process

Modeling Degree-of-Overlap as a Stochastic Process



Modeling Degree-of-Overlap as a Stochastic Process

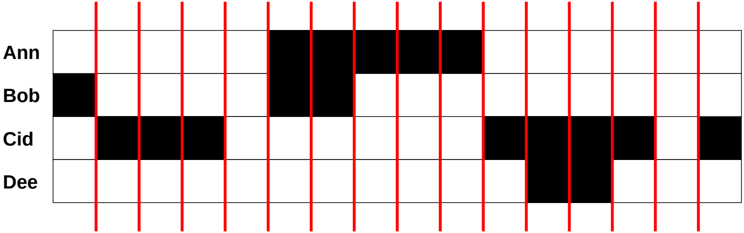
Modeling Degree-of-Overlap as a Stochastic Process

Modeling Degree-of-Overlap as a Stochastic Process

Modeling Degree-of-Overlap as a Stochastic Process

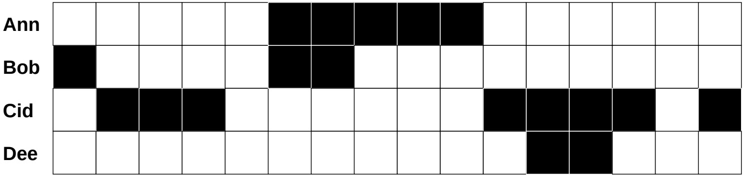
- analysis must be invariant under participant-index rotation

Modeling Degree-of-Overlap as a Stochastic Process



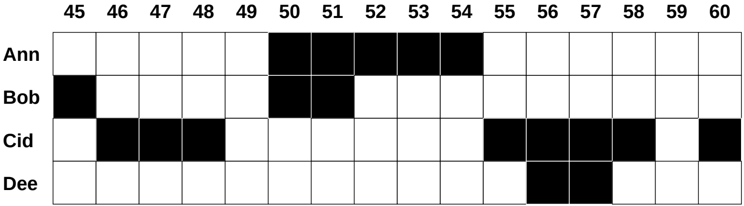
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames

Modeling Degree-of-Overlap as a Stochastic Process



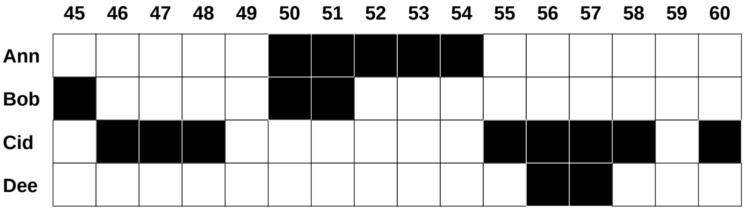
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames

Modeling Degree-of-Overlap as a Stochastic Process



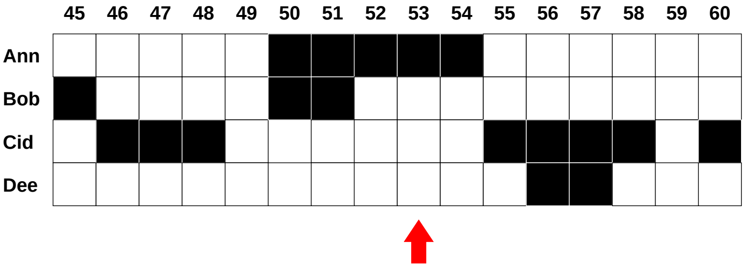
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames

Modeling Degree-of-Overlap as a Stochastic Process



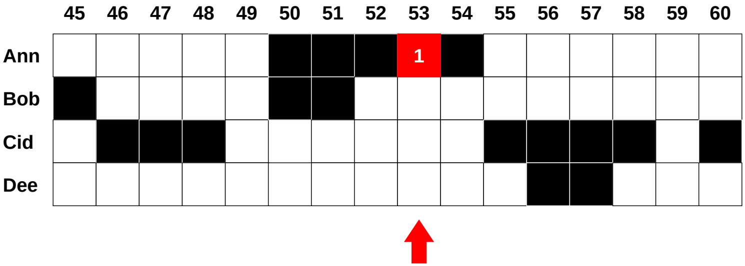
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



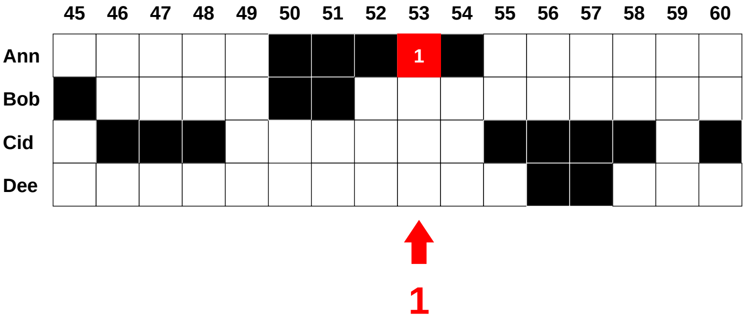
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



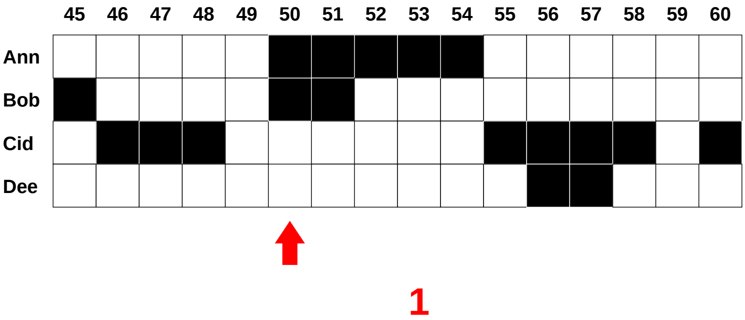
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



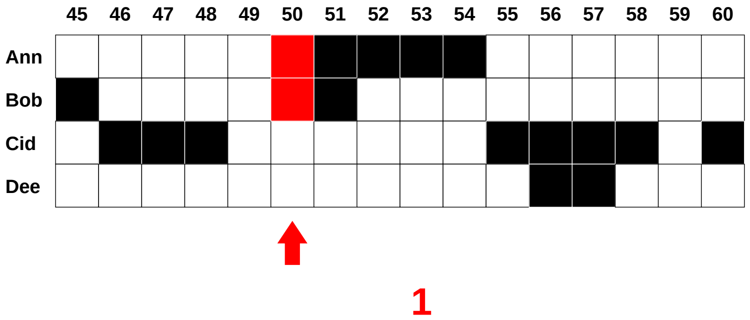
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



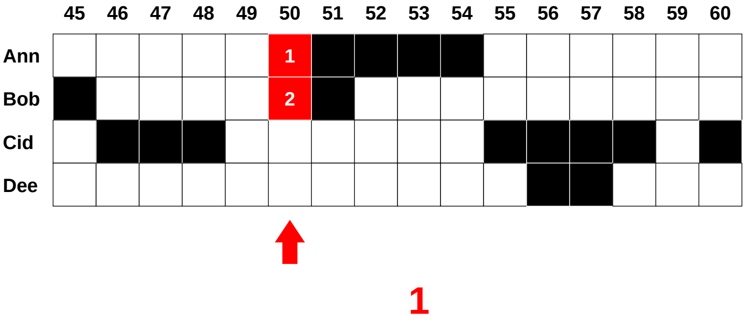
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



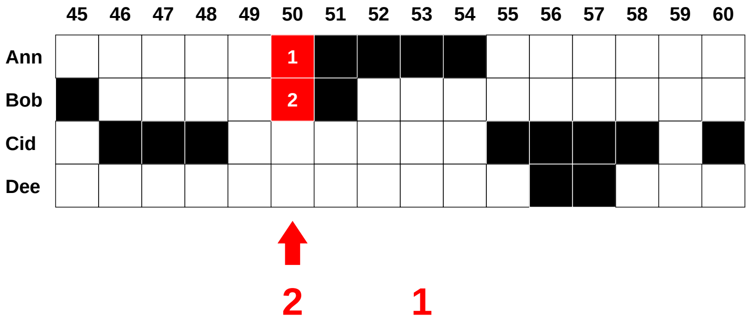
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



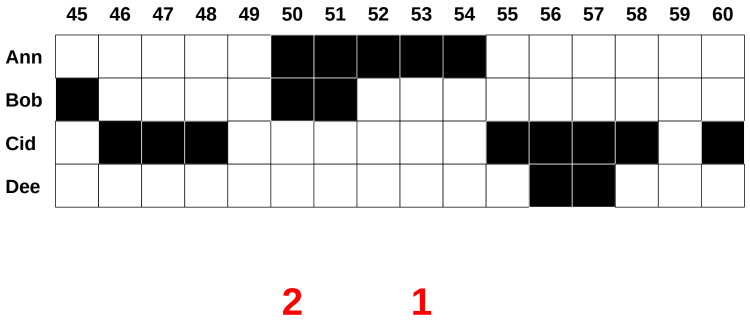
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



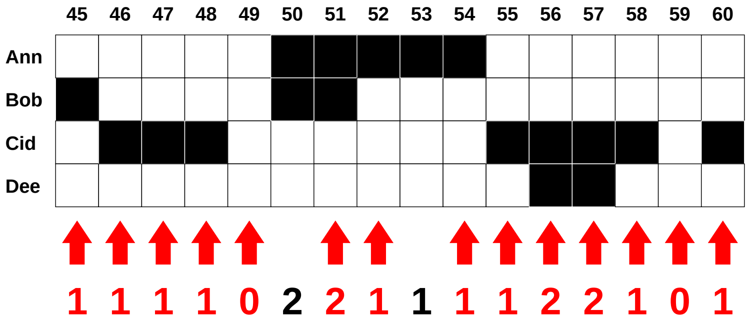
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



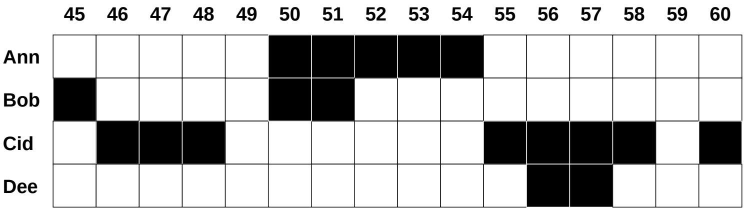
- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

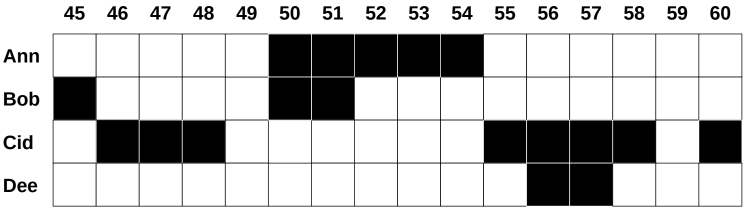
Modeling Degree-of-Overlap as a Stochastic Process



1 1 1 1 0 2 2 1 1 1 1 2 2 1 0 1

- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame

Modeling Degree-of-Overlap as a Stochastic Process



1 1 1 1 0 2 2 1 1 1 1 2 2 1 0 1

- analysis must be invariant under participant-index rotation
- discretize in time using non-overlapping 100-ms frames
- compute the number of speaking participants in each frame
- model integer sequence **using a 1st-order *N*-gram**

Conversational Corpus

Experiments use the **ICSI Meeting Corpus** (Janin et al, 2003):

- ICSI meetings are **conversations**, as per (Sacks et al, 1974)
- **natural**: would have occurred even if were not recorded
- 75 conversations
- each approximately 60 minutes in duration
- each with fixed number of participants, between 3 and 9
- manually transcribed and automatically **forced-aligned**
- manually segmented into dialog acts and **labeled with type** (Shriberg et al, 2004)

Number of Speakers versus Number of Participants

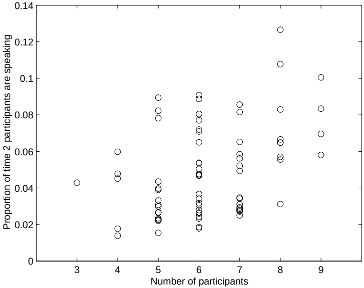
- Note that **the number of participants is different** for different ICSI meetings.

Number of Speakers versus Number of Participants

- Note that **the number of participants is different** for different ICSI meetings.
- But proposing to model degree-of-overlap **unconditioned** on the number of participants, across meetings.
- Is this valid?

Number of Speakers versus Number of Participants

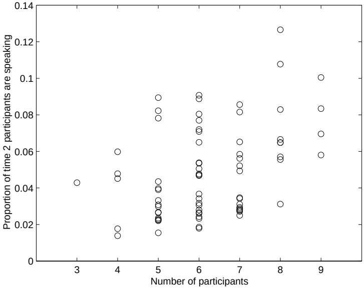
- Note that **the number of participants is different** for different ICSI meetings.
- But proposing to model degree-of-overlap **unconditioned** on the number of participants, across meetings.
- Is this valid?



- correlation between number of meeting attendees and proportion of meeting time during which two attendees speak simultaneously is weak

Number of Speakers versus Number of Participants

- Note that **the number of participants is different** for different ICSI meetings.
- But proposing to model degree-of-overlap **unconditioned** on the number of participants, across meetings.
- Is this valid?



- correlation between number of meeting attendees and proportion of meeting time during which two attendees speak simultaneously is weak
- Pearson's correlation coefficient: 0.411

Is N-Gram Modeling Appropriate?

Yes, provided that:

- The durations of contiguous intervals of same-degree overlap have an exponential distribution.

Is N-Gram Modeling Appropriate?

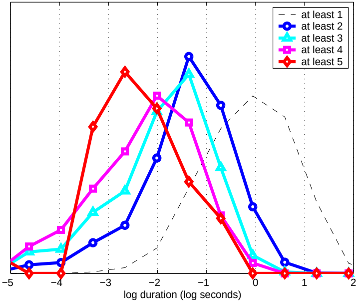
Yes, provided that:

- The durations of contiguous intervals of same-degree overlap have an exponential distribution.
- Approximately: log-normal; at least unimodal.

Is *N*-Gram Modeling Appropriate?

Yes, provided that:

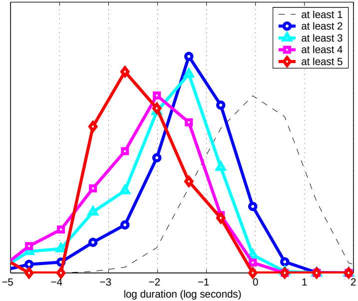
- The durations of contiguous intervals of same-degree overlap have an exponential distribution.
- Approximately: log-normal; at least unimodal.



Is *N*-Gram Modeling Appropriate?

Yes, provided that:

- The durations of contiguous intervals of same-degree overlap have an exponential distribution.
- Approximately: log-normal; at least unimodal.

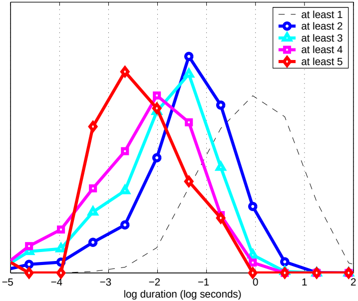


- approximately log-normal as required

Is *N*-Gram Modeling Appropriate?

Yes, provided that:

- The durations of contiguous intervals of same-degree overlap have an exponential distribution.
- Approximately: log-normal; at least unimodal.



- approximately log-normal as required
- also: the lower the degree-of-overlap, the longer the interval

Require **Two** Models

Require **Two** Models

FORWARD IN TIME **BACKWARD IN TIME**

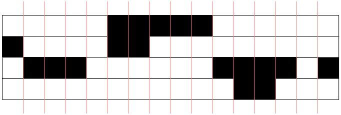
Require **Two** Models

Require Two Models

FORWARD IN TIME

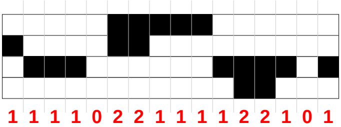
Require **Two** Models

FORWARD IN TIME



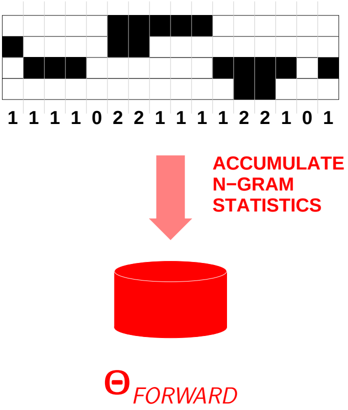
Require Two Models

FORWARD IN TIME

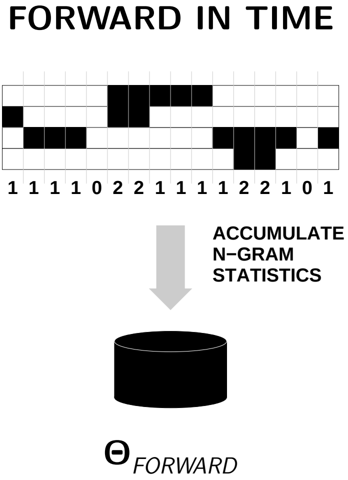


Require **Two** Models

FORWARD IN TIME



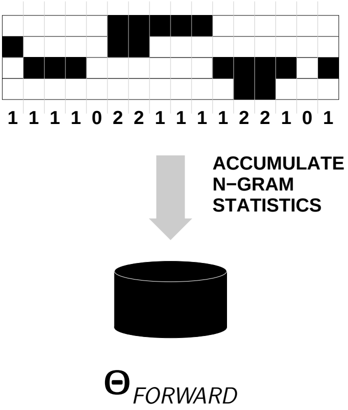
Require Two Models



Require **Two** Models

FORWARD IN TIME

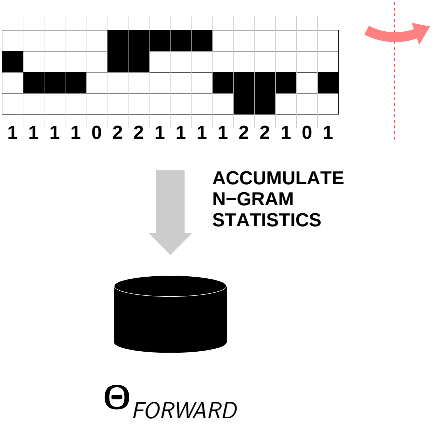
BACKWARD IN TIME



Require Two Models

FORWARD IN TIME

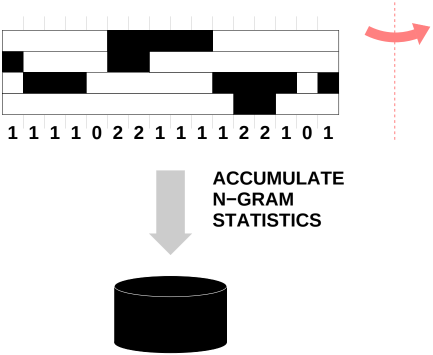
BACKWARD IN TIME



Require Two Models

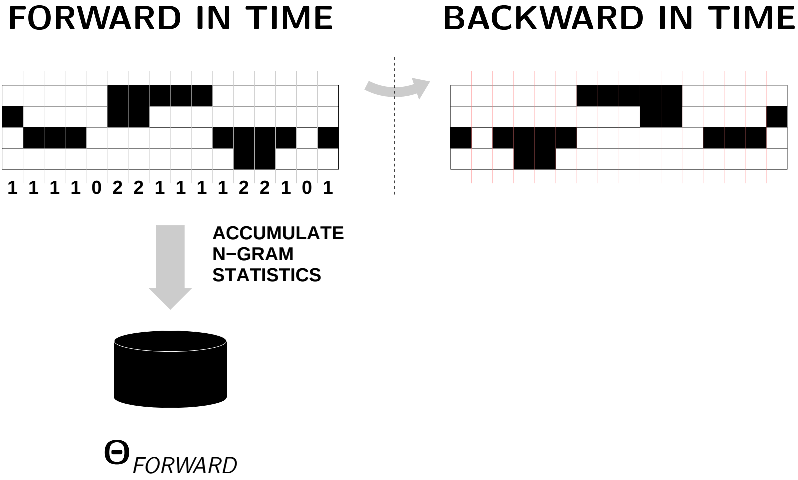
FORWARD IN TIME

BACKWARD IN TIME

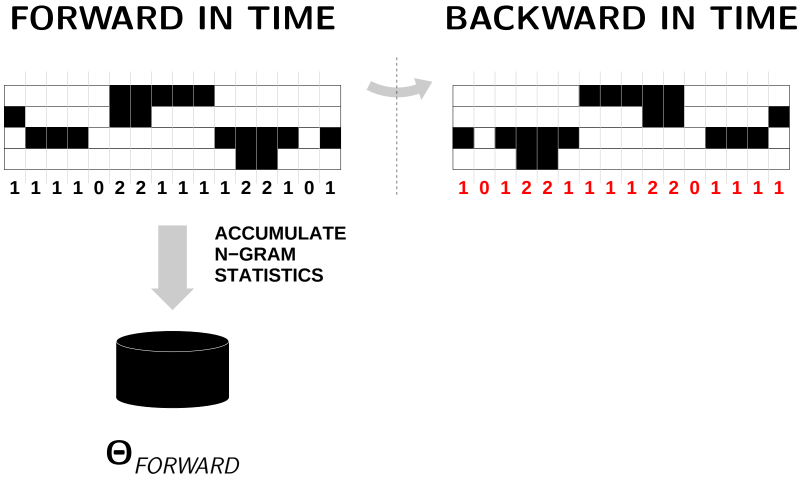


$\Theta_{FORWARD}$

Require Two Models



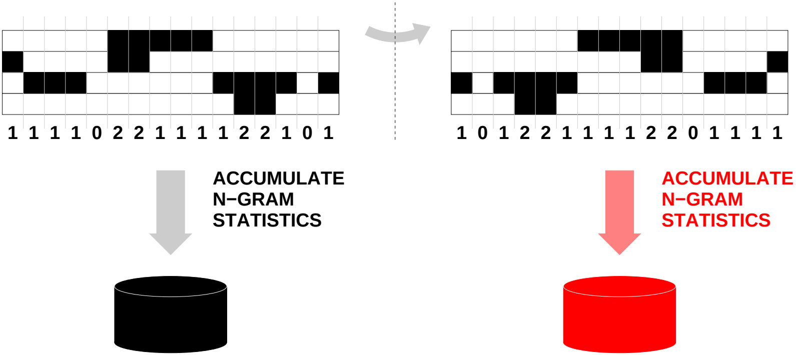
Require Two Models



Require **Two** Models

FORWARD IN TIME

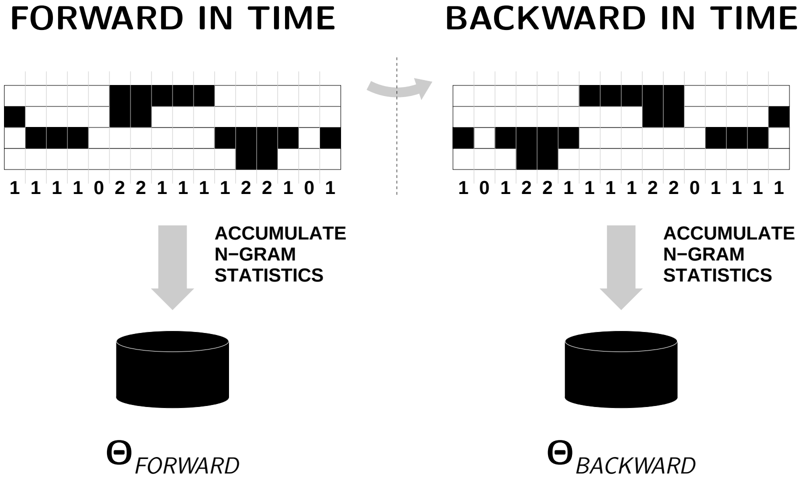
BACKWARD IN TIME



$\Theta_{FORWARD}$

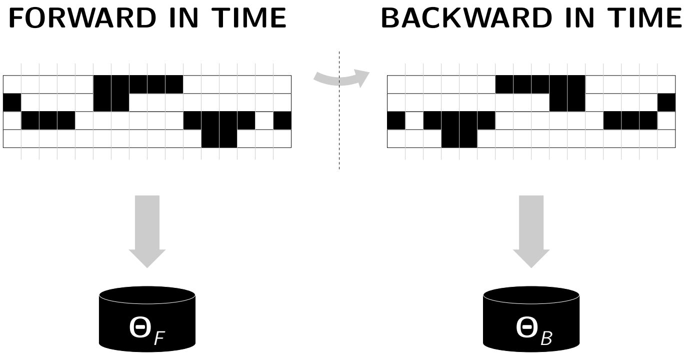
$\Theta_{BACKWARD}$

Require Two Models

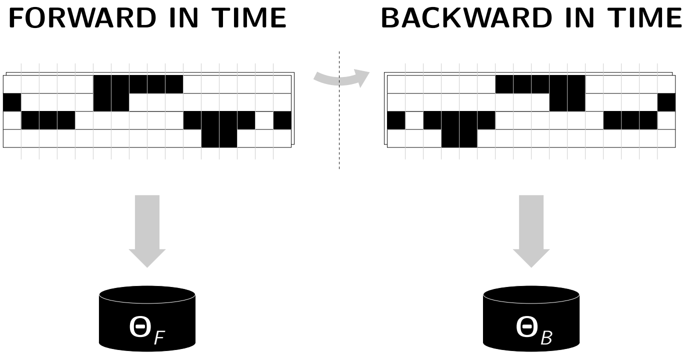


Detecting Time's Arrow in an Unseen Test Chronogram

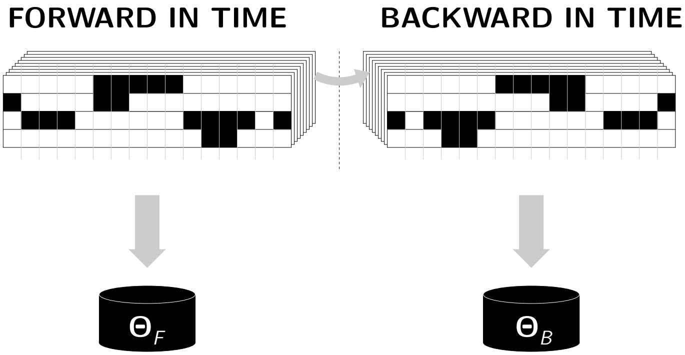
Detecting Time's Arrow in an Unseen Test Chronogram



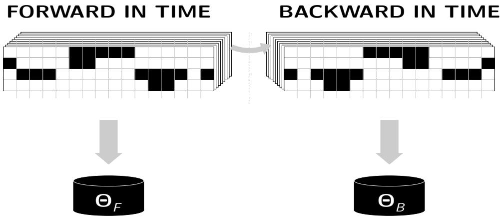
Detecting Time's Arrow in an Unseen Test Chronogram



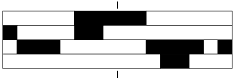
Detecting Time's Arrow in an Unseen Test Chronogram



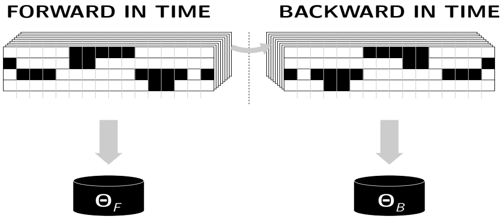
Detecting Time's Arrow in an Unseen Test Chronogram



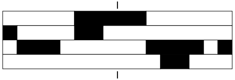
1. Take test chronogram **Q**.



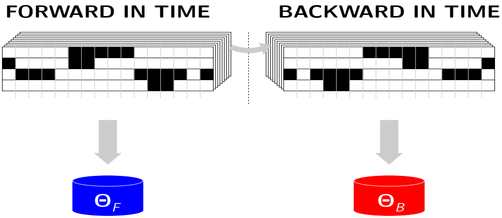
Detecting Time's Arrow in an Unseen Test Chronogram



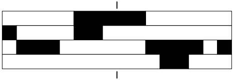
1. Take test chronogram $\mathbf{Q}$.
2. Pick **random** direction $d \in \{F, B\}$: $\mathbf{Q} \mapsto \mathbf{Q}'$.



Detecting Time's Arrow in an Unseen Test Chronogram

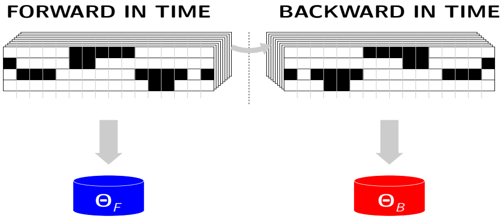


1. Take test chronogram $\mathbf{Q}$.
2. Pick **random** direction $d \in \{F, B\}$: $\mathbf{Q} \mapsto \mathbf{Q}'$.

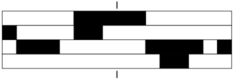


3. Compute $P\left(\mathbf{Q}' | \boldsymbol{\Theta}_F\right)$ and $P\left(\mathbf{Q}' | \boldsymbol{\Theta}_B\right)$.

Detecting Time's Arrow in an Unseen Test Chronogram



1. Take test chronogram $\mathbf{Q}$.
2. Pick **random** direction $d \in \{F, B\}$: $\mathbf{Q} \mapsto \mathbf{Q}'$.



3. Compute $P\left(\mathbf{Q}' | \mathbf{\Theta}_F\right)$ and $P\left(\mathbf{Q}' | \mathbf{\Theta}_B\right)$.
4. Guess $\hat{d}$ yielding higher likelihood.

Classification in a Round Robin Evaluation

- Propose to **not** assess statistical significance.

Classification in a Round Robin Evaluation

- Propose to **not** assess statistical significance.
- Instead, assess how well can **classify** the direction d ...
  - ... in conversations **unseen during training**
  - a more stringent requirement for discarding null hypothesis

Classification in a Round Robin Evaluation

- Propose to **not** assess statistical significance.
- Instead, assess how well can **classify** the direction *d* ...
  - ... in conversations **unseen during training**
  - a more stringent requirement for discarding null hypothesis
- ICSI Corpus contains 75 conversations:

Classification in a Round Robin Evaluation

- Propose to **not** assess statistical significance.
- Instead, assess how well can **classify** the direction d ...
  - ... in conversations **unseen during training**
  - a more stringent requirement for discarding null hypothesis

- ICSI Corpus contains 75 conversations:
  1. Pick each conversation as the **test** conversation $\mathbf{Q}$.
  2. Train $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ on remaining 74 conversations.
  3. Pick direction d (50%/50%), form $d : \mathbf{Q} \mapsto \mathbf{Q}'$.
  4. Infer direction $\hat{d} \doteq \arg\max_d P\left(\mathbf{Q}' | \mathbf{\Theta}_d\right)$.

Classification in a Round Robin Evaluation

- Propose to **not** assess statistical significance.
- Instead, assess how well can **classify** the direction d ...
  - ... in conversations **unseen during training**
  - a more stringent requirement for discarding null hypothesis

- ICSI Corpus contains 75 conversations:
  1. Pick each conversation as the **test** conversation $\mathbf{Q}$.
  2. Train $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ on remaining 74 conversations.
  3. Pick direction d (50%/50%), form $d : \mathbf{Q} \mapsto \mathbf{Q}'$.
  4. Infer direction $\hat{d} \doteq \arg\max_d P\left(\mathbf{Q}' | \mathbf{\Theta}_d\right)$.

- Accuracy (A): count how often $\hat{d} = d$, divide by 75.

Classification in a Round Robin Evaluation

- Propose to **not** assess statistical significance.
- Instead, assess how well can **classify** the direction d ...
  - ... in conversations **unseen during training**
  - a more stringent requirement for discarding null hypothesis
- ICSI Corpus contains 75 conversations:
  1. Pick each conversation as the **test** conversation $\mathbf{Q}$.
  2. Train $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ on remaining 74 conversations.
  3. Pick direction d (50%/50%), form $d : \mathbf{Q} \mapsto \mathbf{Q}'$.
  4. Infer direction $\hat{d} \doteq \arg\max_d P\left(\mathbf{Q}' | \mathbf{\Theta}_d\right)$.
- Accuracy (A): count how often $\hat{d} = d$, divide by 75.
  - If chronograms are symmetric in time, expect $[A] = 50\%$.

Classification in a Round Robin Evaluation

- Propose to **not** assess statistical significance.
- Instead, assess how well can **classify** the direction d ...
  - ... in conversations **unseen during training**
  - a more stringent requirement for discarding null hypothesis

- ICSI Corpus contains 75 conversations:
  1. Pick each conversation as the **test** conversation $\mathbf{Q}$.
  2. Train $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ on remaining 74 conversations.
  3. Pick direction d (50%/50%), form $d : \mathbf{Q} \mapsto \mathbf{Q}'$.
  4. Infer direction $\hat{d} \doteq \arg\max_d P\left(\mathbf{Q}'|\mathbf{\Theta}_d\right)$.

- Accuracy (A): count how often $\hat{d} = d$, divide by 75.
  - If chronograms are symmetric in time, expect $[A] = 50\%$.
- Chance-corrected accuracy,

$$ccA \doteq \frac{A - [A]}{1 - [A]}$$

Numerical Results

- Over all 75 conversations, $A = 99$**% and** $ccA = 97$**%**
- Can comfortably discard the null hypothesis H_0, that chronograms are left-right symmetric.

Numerical Results

- Over all 75 conversations, $A = 99$**% and** $ccA = 97$**%**
- Can comfortably discard the null hypothesis H_0, that chronograms are left-right symmetric.

- Note that temporal asymmetry **must** be due to overlap.

The Link Between Asymmetry and Overlap

- Consider a sequence consisting exclusively of 0s and 1s:

  e.g., $\ldots 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, \ldots$

The Link Between Asymmetry and Overlap

- Consider a sequence consisting exclusively of 0s and 1s:

  e.g., $\ldots 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, \ldots$

1. The number of $\{0 \rightarrow 1\}$ and $\{1 \rightarrow 0\}$ transitions is equal.

The Link Between Asymmetry and Overlap

- Consider a sequence consisting exclusively of 0s and 1s:

$$\text{e.g., } \ldots 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, \ldots$$

1. The number of $\{0 \rightarrow 1\}$ and $\{1 \rightarrow 0\}$ transitions is equal.
2. The models $\boldsymbol{\Theta}_F$ and $\boldsymbol{\Theta}_B$ are equal.

The Link Between Asymmetry and Overlap

- Consider a sequence consisting exclusively of 0s and 1s:

  e.g., $\ldots 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, \ldots$

  1. The number of $\{0 \rightarrow 1\}$ and $\{1 \rightarrow 0\}$ transitions is equal.
  2. The models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ are equal.
  3. Cannot discriminate between F and B directions.

The Link Between Asymmetry and Overlap

- Consider a sequence consisting exclusively of 0s and 1s:

  e.g., $\ldots 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, \ldots$

  1. The number of $\{0 \rightarrow 1\}$ and $\{1 \rightarrow 0\}$ transitions is equal.
  2. The models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ are equal.
  3. Cannot discriminate between F and B directions.

- Asymmetry in a chronogram is:

  1. impossible if, $\forall t$, the degree-of-overlap is ≤ 1;

The Link Between Asymmetry and Overlap

- Consider a sequence consisting exclusively of 0s and 1s:

  e.g., $\ldots 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, \ldots$

  1. The number of $\{0 \rightarrow 1\}$ and $\{1 \rightarrow 0\}$ transitions is equal.
  2. The models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ are equal.
  3. Cannot discriminate between F and B directions.

- Asymmetry in a chronogram is:

  1. impossible if, $\forall t$, the degree-of-overlap is ≤ 1;
  2. impossible if, $\forall t$, the degree-of-overlap changes by ≤ 1;

The Link Between Asymmetry and Overlap

- Consider a sequence consisting exclusively of 0s and 1s:

  e.g., $\ldots 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, \ldots$

  1. The number of $\{0 \rightarrow 1\}$ and $\{1 \rightarrow 0\}$ transitions is equal.
  2. The models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ are equal.
  3. Cannot discriminate between F and B directions.

- Asymmetry in a chronogram is:

  1. impossible if, $\forall t$, the degree-of-overlap is ≤ 1;
  2. impossible if, $\forall t$, the degree-of-overlap changes by ≤ 1;
  3. possible (but not guaranteed) if $\exists t$ at which

     the degree-of-overlap changes by ≥ 2.

More Necessary and Sufficient Conditions for Asymmetry

- Models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ can be inspected directly.

More Necessary and Sufficient Conditions for Asymmetry

- Models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ can be inspected directly.
- In a zero-bounded sequence, a degree-of-overlap of 2 occurs:
    1. most often: (0...0)(1...1)(2...2)(1...1)(0...0)

More Necessary and Sufficient Conditions for Asymmetry

- Models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ can be inspected directly.
- In a zero-bounded sequence, a degree-of-overlap of 2 occurs:
    1. most often: (0...0)(1...1)(2...2)(1...1)(0...0)
    2. less often: (0...0)(2...2)(1...1)(0...0)

More Necessary and Sufficient Conditions for Asymmetry

- Models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ can be inspected directly.
- In a zero-bounded sequence, a degree-of-overlap of 2 occurs:
  1. most often: (0...0)(1...1)(2...2)(1...1)(0...0)
  2. less often: (0...0)(2...2)(1...1)(0...0)
  3. least often: (0...0)(1...1)(2...2)(0...0)

More Necessary and Sufficient Conditions for Asymmetry

- Models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ can be inspected directly.
- In a zero-bounded sequence, a degree-of-overlap of 2 occurs:
  1. most often: (0...0)(1...1)(2...2)(1...1)(0...0)
  2. less often: (0...0)(2...2)(1...1)(0...0)
  3. least often: (0...0)(1...1)(2...2)(0...0)
- Case 1 is 1st-order-Markov-symmetric.
  - Cannot account for the left-to-right asymmetry in chronograms.

More Necessary and Sufficient Conditions for Asymmetry

- Models $\mathbf{\Theta}_F$ and $\mathbf{\Theta}_B$ can be inspected directly.
- In a zero-bounded sequence, a degree-of-overlap of 2 occurs:
  1. most often: (0...0)(1...1)(2...2)(1...1)(0...0)
  2. less often: (0...0)(2...2)(1...1)(0...0)
  3. least often: (0...0)(1...1)(2...2)(0...0)
- Case 1 is 1st-order-Markov-symmetric.
  - Cannot account for the left-to-right asymmetry in chronograms.
- Time's arrow is discernable in chronograms largely because case 2 and case 3 occur with **unequal** frequency.

Finding the Culprits

- Would like to know what **kinds of speech phenomena** lead to more $\{0 \rightarrow 2\}$ transitions than to $\{2 \rightarrow 0\}$ transitions.

Finding the Culprits

- Would like to know what **kinds of speech phenomena** lead to more $\{0 \rightarrow 2\}$ transitions than to $\{2 \rightarrow 0\}$ transitions.
- Propose to investigate (content-neutral) **dialog act (DA)** types as a subclassification of all speech.

Finding the Culprits

- Would like to know what **kinds of speech phenomena** lead to more $\{0 \rightarrow 2\}$ transitions than to $\{2 \rightarrow 0\}$ transitions.
- Propose to investigate (content-neutral) **dialog act (DA)** types as a subclassification of all speech.

- The ICSI Corpus is annotated with a rich tagset, including:
  - unlabeled $\mathcal{X}$: not speech, undecipherable, undecidable
  - disrupted $\mathcal{D}$: abandoned, interrupted
  - backchannels $\mathcal{B}$: backchannels, assessments, acknowledgments
  - floor mechanisms $\mathcal{F}$: floor grabbers, floor holders, holds
  - propositional: statements, questions

Ablation of Specific-DA Deployment

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.
- To test the impact of DA type $\mathcal{T}$ on asymmetry:
  1. Compute *ccA* using round robin paradigm, as in Part I.

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.
- To test the impact of DA type $\mathcal{T}$ on asymmetry:
  1. Compute ccA using round robin paradigm, as in Part I.
  2. Remove all speech of type $\mathcal{T}$ from the training material.

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.
- To test the impact of DA type $\mathcal{T}$ on asymmetry:
  1. Compute *ccA* using round robin paradigm, as in Part I.
  2. Remove all speech of type $\mathcal{T}$ from the training material.

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.
- To test the impact of DA type $\mathcal{T}$ on asymmetry:
  1. Compute *ccA* using round robin paradigm, as in Part I.
  2. Remove all speech of type $\mathcal{T}$ from the training material.

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.
- To test the impact of DA type $\mathcal{T}$ on asymmetry:
  1. Compute *ccA* using round robin paradigm, as in Part I.
  2. Remove all speech of type $\mathcal{T}$ from the training material.

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.
- To test the impact of DA type $\mathcal{T}$ on asymmetry:
  1. Compute *ccA* using round robin paradigm, as in Part I.
  2. Remove all speech of type $\mathcal{T}$ from the training material.

  3. Remove all speech of type $\mathcal{T}$ from the test chronogram.

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.
- To test the impact of DA type $\mathcal{T}$ on asymmetry:
  1. Compute ccA using round robin paradigm, as in Part I.
  2. Remove all speech of type $\mathcal{T}$ from the training material.

  3. Remove all speech of type $\mathcal{T}$ from the test chronogram.
  4. Compute $ccA_{\mathcal{T}}$ using round robin paradigm.

Ablation of Specific-DA Deployment

- Re-use the experimental methodology of Part I.
- To test the impact of DA type $\mathcal{T}$ on asymmetry:
  1. Compute ccA using round robin paradigm, as in Part I.
  2. Remove all speech of type $\mathcal{T}$ from the training material.

  3. Remove all speech of type $\mathcal{T}$ from the test chronogram.
  4. Compute $ccA_{\mathcal{T}}$ using round robin paradigm.
  5. Compare ccA and $ccA_{\mathcal{T}}$.

Results

DA Types Removed	Duration of Speech Remaining (hh:mm)	*ccA* (%)
none	66:34	97
unlabeled $\mathcal{X}$	63:37 (95.6%)	97
$\mathcal{X}$ ∪ disrupted $\mathcal{D}$	56:44 (85.2%)	89
$\mathcal{X}$ ∪ backchannels $\mathcal{B}$	59:08 (88.8%)	79
$\mathcal{X} \cup \mathcal{D} \cup \mathcal{B}$	52:22 (78.7%)	65
$\mathcal{X}$ ∪ floor mechanisms $\mathcal{F}$	57:03 (85.7%)	89
$\mathcal{X} \cup \mathcal{D} \cup \mathcal{F}$	50:48 (76.3%)	76
$\mathcal{X} \cup \mathcal{D} \cup \mathcal{B} \cup \mathcal{F}$	46:31 (69.9%)	30

Time's arrow can be inferred from chronograms primarily due to:

- disrupted (abandoned or interrupted) DAs, and
- DAs not implementing propositional content.

Conclusions

1. Speech/non-speech chronograms are **asymmetric in time**.

Conclusions

1. Speech/non-speech chronograms are **asymmetric in time**.
2. The asymmetry is due to **entry into** and **egress from** overlap.

Conclusions

1. Speech/non-speech chronograms are **asymmetric in time**.
2. The asymmetry is due to **entry into** and **egress from** overlap.

most common
(undiscriminative)

Conclusions

1. Speech/non-speech chronograms are **asymmetric in time**.
2. The asymmetry is due to **entry into** and **egress from** overlap.

less common

most common
(undiscriminative)

Conclusions

1. Speech/non-speech chronograms are **asymmetric in time**.
2. The asymmetry is due to **entry into** and **egress from** overlap.

less common
(discriminative)

most common
(undiscriminative)

least common
(discriminative)

Conclusions

1. Speech/non-speech chronograms are **asymmetric in time**.
2. The asymmetry is due to **entry into** and **egress from** overlap.

less common
(discriminative)

most common
(undiscriminative)

least common
(discriminative)

3. People are more likely to simultaneously start simultaneous speech than to simultaneously stop simultaneous speech.

Conclusions

1. Speech/non-speech chronograms are **asymmetric in time**.
2. The asymmetry is due to **entry into** and **egress from** overlap.

less common
(discriminative)

most common
(undiscriminative)

least common
(discriminative)

3. People are more likely to simultaneously start simultaneous speech than to simultaneously stop simultaneous speech.
4. Speech to which this pertains is found in dialog acts:
   - which are **not** successfully brought to completion, or
   - whose pragmatic function is **not** information exchange.

Potential Impact

- Theoretical:

Potential Impact

- Theoretical:
  - Can empirically validate on large data the claims of conversation analysis.
    - E.g., *"Talk by MORE than two at a time seems to be reduced to two (or to one) even more effectively than talk by two is reduced to one"* (Schegloff, 2000).

Potential Impact

- Theoretical:
  - Can empirically validate on large data the claims of conversation analysis.
    - E.g., *"Talk by MORE than two at a time seems to be reduced to two (or to one) even more effectively than talk by two is reduced to one"* (Schegloff, 2000).
    - See paper for many examples.

Potential Impact

- Theoretical:
  - Can empirically validate on large data the claims of conversation analysis.
    - E.g., *"Talk by MORE than two at a time seems to be reduced to two (or to one) even more effectively than talk by two is reduced to one"* (Schegloff, 2000).
    - See paper for many examples.
  - One step along the way to proposing an ecological theory of pragmatic (non-propositional-content) function in multi-party conversation, and its relationship to cognitive load.

Potential Impact

- Theoretical:
  - Can empirically validate on large data the claims of conversation analysis.
    - E.g., *"Talk by MORE than two at a time seems to be reduced to two (or to one) even more effectively than talk by two is reduced to one"* (Schegloff, 2000).
    - See paper for many examples.
  - One step along the way to proposing an ecological theory of pragmatic (non-propositional-content) function in multi-party conversation, and its relationship to cognitive load.
- Technological:

Potential Impact

- Theoretical:
  - Can empirically validate on large data the claims of conversation analysis.
    - E.g., *"Talk by MORE than two at a time seems to be reduced to two (or to one) even more effectively than talk by two is reduced to one"* (Schegloff, 2000).
    - See paper for many examples.
  - One step along the way to proposing an ecological theory of pragmatic (non-propositional-content) function in multi-party conversation, and its relationship to cognitive load.
- Technological:
  - Construction of prior probability models for speech activity detection in multi-party conversations.
    - E.g., constrain hypothesized transitions into and out of overlap intervals.

THANK YOU

The work was supported in part by the Riksbankens Jubileumsfond (RJ) project Samtalets Prosodi.