Dialog Formation from the Extracted Clusters of Dendrograms
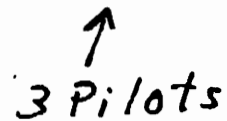
&

Perfomance Measurment

Clustering Information

- Clustering is based on similarity of transmission pairs:

    - speaker/channel characteristics

    - recognized word or phoneme sequence

- Agglomerative clustering is performed to create tree-structured groupings ("dendrograms")

Controller-Pilot Dendogram Based on Acoustics

423 Utterances

3 Pilots

3 Controllers in with Pilots

Distance Measure Between two Transmissions

- Symmetric
  - No training and test

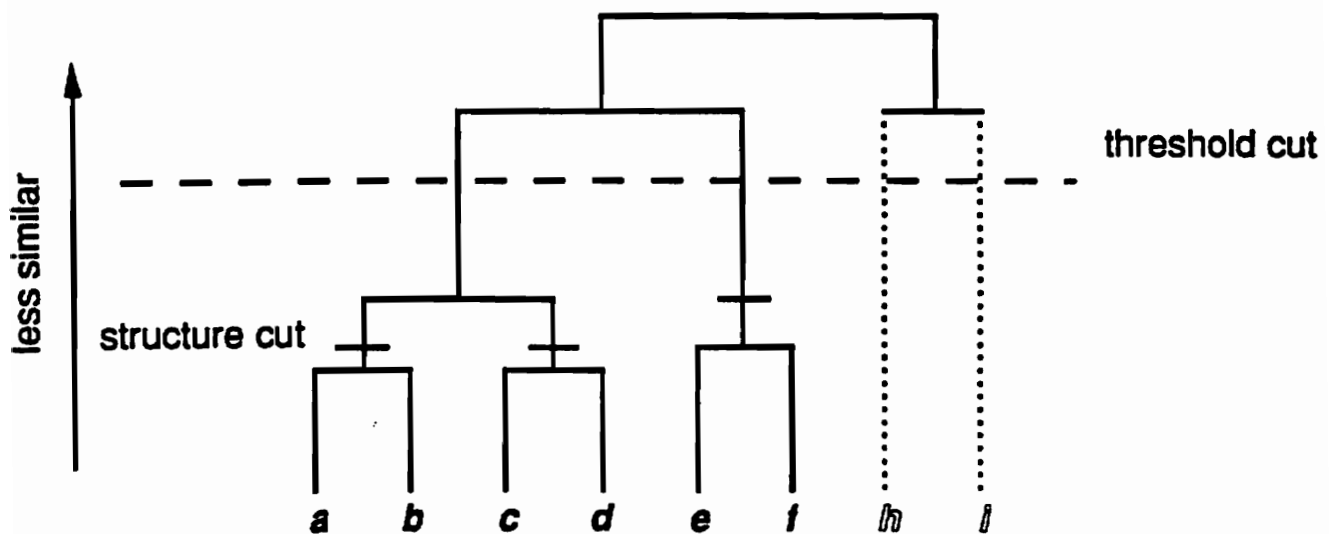- Depends on number of samples in each of the transmissions

Current Distance Measure $dist = d_{COV} + d_{MEAN}$

$$d_{COV} = \frac{N}{2} log|W| - \frac{N_1}{2} log|S_1| - \frac{N_2}{2} log|S_2|$$

$$d_{MEAN} = \frac{N}{2} log[1 + \frac{N_1 N_2}{N}(\bar{x}_1 - \bar{x}_2)'W^{-1}(\bar{x}_1 - \bar{x}_2)]$$

Dendrogram-Based Clustering

- Cutting a dendrogram forms a partition of the data

- Larger clusters are generally less "pure"

- Cuts can be based on threshold on similarity or on structure of the dendrogram

Combining Clusters

- Problem: several separate sources of information are available for dialog formation. How can they be combined?

- Goal: "pure" clusters with high confidence of containing only one pilot or controller

- Approach: form clusters independently from separate sources and the search for consistency

Combining Information is Generic
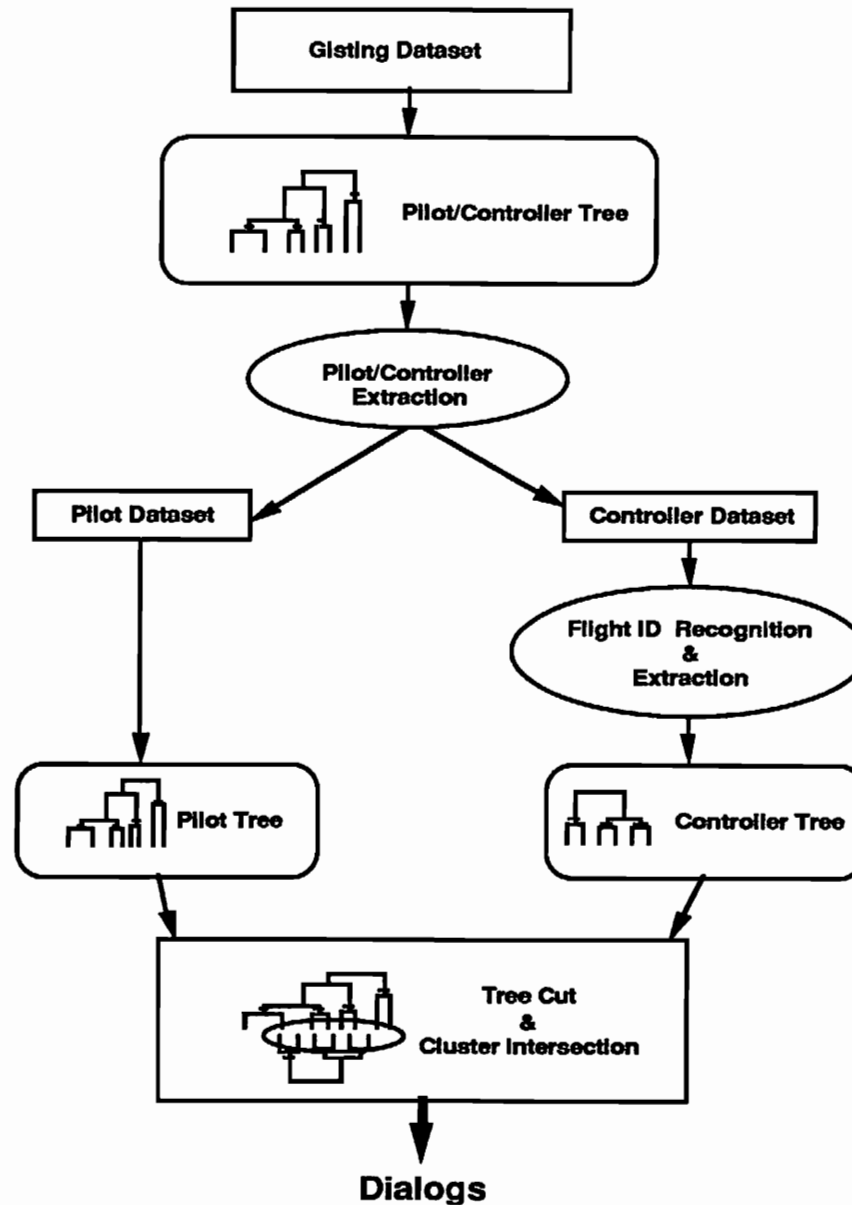
Consistency of Clusters

Pilot and controller transmissions from a single dialog are almost always adjacent

Forming Consistent Clusters

- Cut pilot and controller dendrograms to get almost-single-speaker clusters

    - structural criterion (e.g., 2 merges from bottom of tree)

    - distance-based cut (e.g., mean distance at top of tree)

- Cuts are equivalent to partition of pilots into sets $P_1, \ldots, P_M$ and controllers into sets $C_1, \ldots, C_N$.
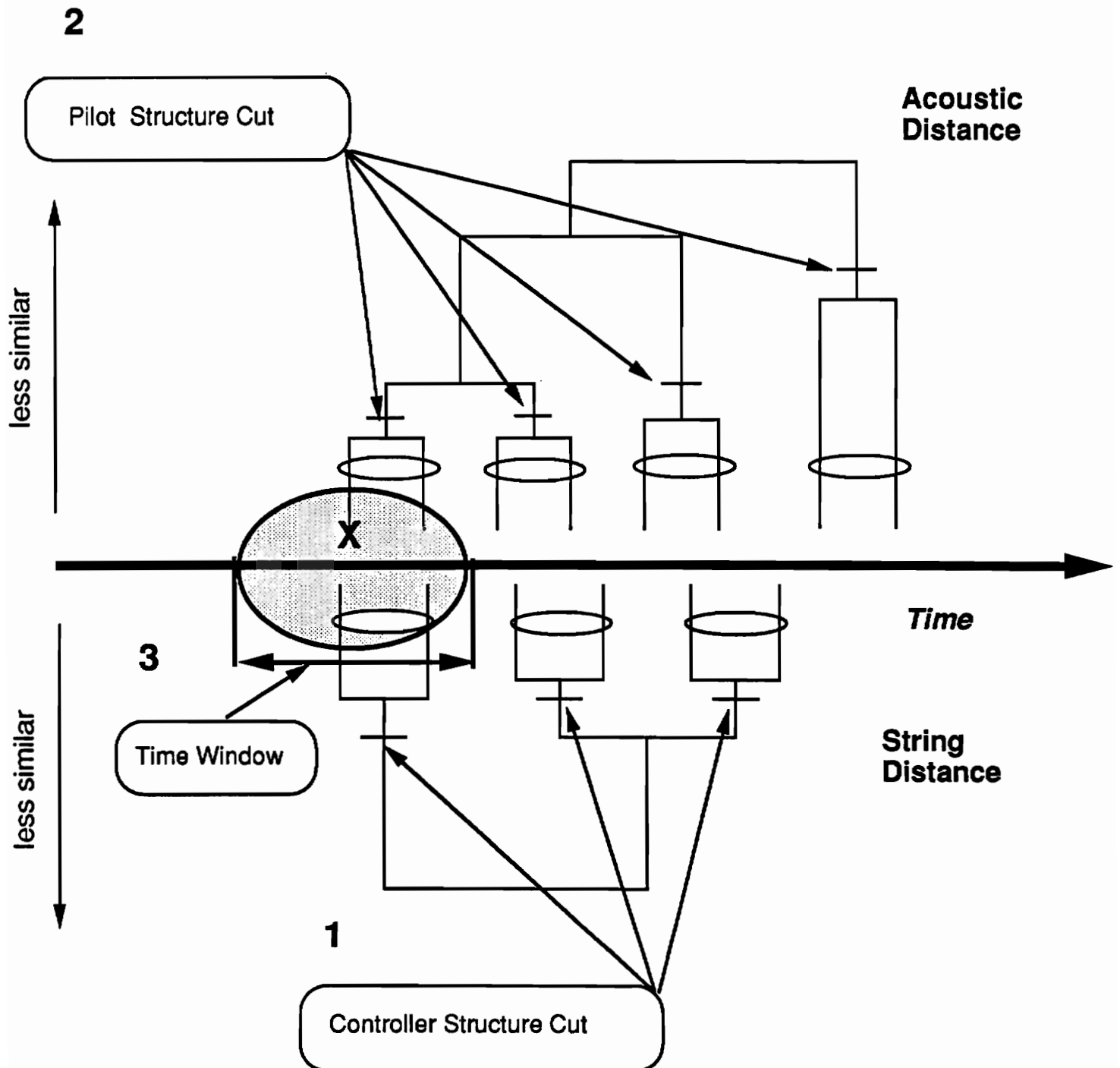
Intersecting Clusters

- For each C_i

    for each P_j

    $\tilde{C}_{ij} = \{c | c \in C_i, c \text{ is adjacent to some } p \in P_j\}$
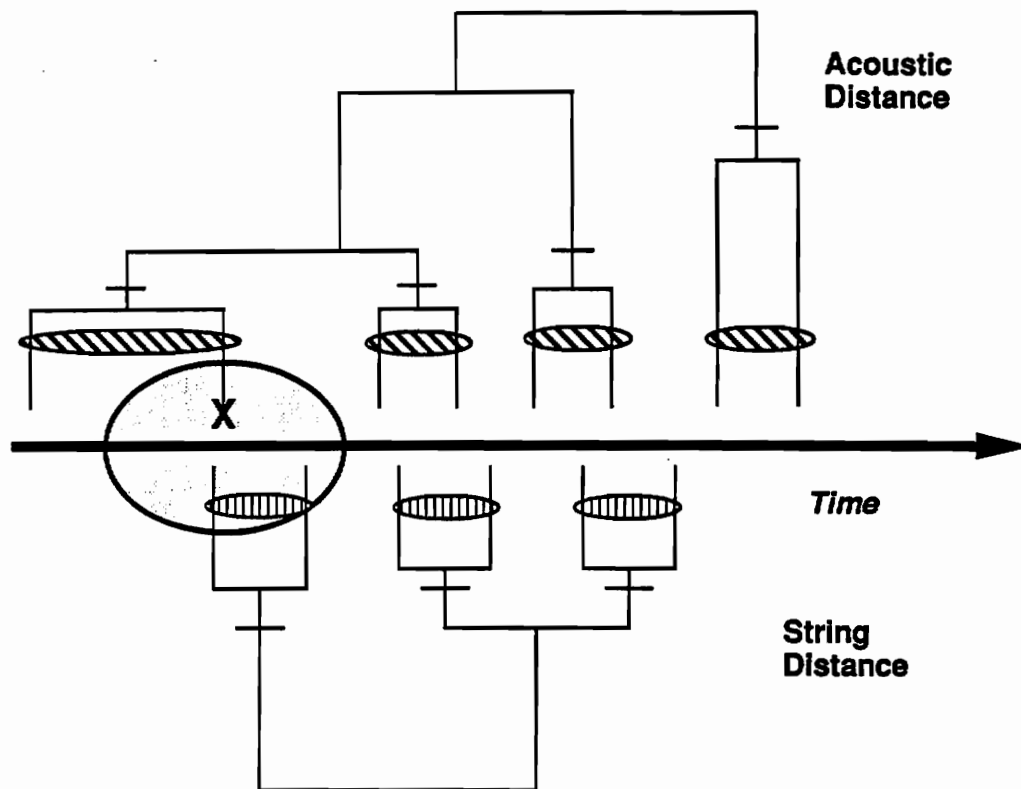
- Let $C'_1, \ldots, C'_{N'}$ be the sets $\tilde{C}_{ij}$ with more than one element where no set is a subset of another

Three Parameters
for Forming Dialogs
from Clusters

2

Pilot Structure Cut

Acoustic
Distance

less similar

X

Time

3

Time Window

less similar

String
Distance

1

Controller Structure Cut

Performance Measurement

How Dialogs Are Arrayed in Time

Associated Dialogs for Pilot/Controller Cluster 2-5-3, Tape14_D1N