

# *Analysis and Visualization of Index Words from Audio Transcripts of Instructional Videos*

Alexander Haubold and John R. Kender  
Department of Computer Science  
Columbia University, New York  
{ahaubold,jrk}@cs.columbia.edu

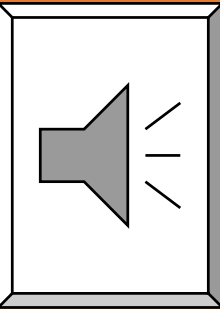
# Overview

- Motivation
- Transcript Generation
- Target Corpus
- Filtering *Index Phrases*
- Filtering *Word Pairs*
- Visualization of Textual Information
- Future Investigations

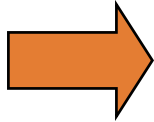
# Motivation

- Automatic Speech Recognition (ASR)  
Transcripts typically used for searching, categorization of video databases
- Lecture videos = dozens of contextually connected entities
- Typical course: 10 to 30 lectures (70 or 120 min), lecture: 5k – 14k words  $\approx$  150k words
- Need indices across lectures, courses
- Extract and display structure of entire course using key words/phrases

# Process Overview

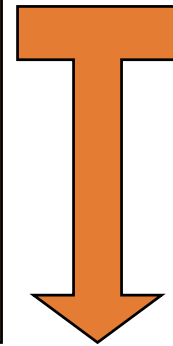


Lecture Audio



... aum sell and it is its structure doesn't provide a way to find something like a binary tree provides a way of looking for 27 and by treat it is given a pointer to the router the treehouse where it ought and emulate ...

Imperfect Transcript

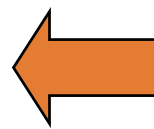


Big Endian Binary search tree  
...  
Pointers

Textbook Index or Manual Index

Chapter Transcript Match		Transcript Index Map									Lecture Title
	ALL	1	2	3	4	5	6	7	8	9	1
01.stein.txt	87...	48...	96...	49...	55...	47...	29...	45...	50...	52...	29
Transcript Index Map	Lecture Transcript Similarity										
05.stei	06.stei	07.stei	08.stei	09.stei	10.stei	11.stei	12.stei	13.stei			
									62...	44...	40...
									32...	55...	31...
									64...	12...	80...
									65...	54...	59...
									45...	50...	32...
									12...	61...	77...
									46...	61...	37...
									39...	41...	27...
									37...	58...	45...
									29...	24...	25...
									15...	20...	26...

Interactive Visualization



...  
Binary  
Binary tree  
...  
Pointer  
Structure  
Tree  
...

Index Words and Phrases

...  
find binary  
find structure  
find tree  
binary tree  
binary pointer  
binary structure  
tree pointer  
tree structure  
...

Word Pairs

# *Transcript Generation*

- ASR transcript from IBM<sup>®</sup> ViaVoice<sup>®</sup>
- Highly compressed lecture video
  - Poor audio quality
  - Word Error Rate of 75%
- Training: little (3%) overall improvement
  - Number of unique index phrases  $\approx$  same
  - But: difference in identified index phrases 20%
  - Best to combine trained and untrained results

# Target Corpus

- Lectures: rich in subject-specific terms
- Define:
  - *Theme phrases*: General tenor for contents of course
    - in many ( $> \frac{1}{4}$ ) transcripts
  - *Topic phrase*: Highlight specific topics for lectures
    - in few ( $< \frac{1}{4}$ ) transcripts
  - *Illustration phrases*: unique terms for examples
    - Hard to identify in highly imperfect transcripts

# Filtering Index Phrases

- Structured approach
- Use corpus of expected phrases: index of course textbook
  - Capture key phrases of length 1-3
  - Rarely longer; index reflects likelihood
  - (1) Collapse indentation hierarchy
  - (2) Remove stop words in beginning & end of each line
  - (3) Stem

amortized analysis	(3)
accounting method of	(1,2,3)
aggregate analysis of	(1,2)
...	
call by value	

# Filtering Word Pairs

- Unstructured approach
  - Address: speech in lecture fragmented
- Use textbook index to filter keywords
  - Remove structure from index: use only words
  - Word pairs in transcript = index words separated by  $\leq 10$  words

multiple instruction	call structural
multiple operation	call hazard
multiple very	call instruction
multiple word	call compaction
multiple processor	call step



# *Results of Filtering*

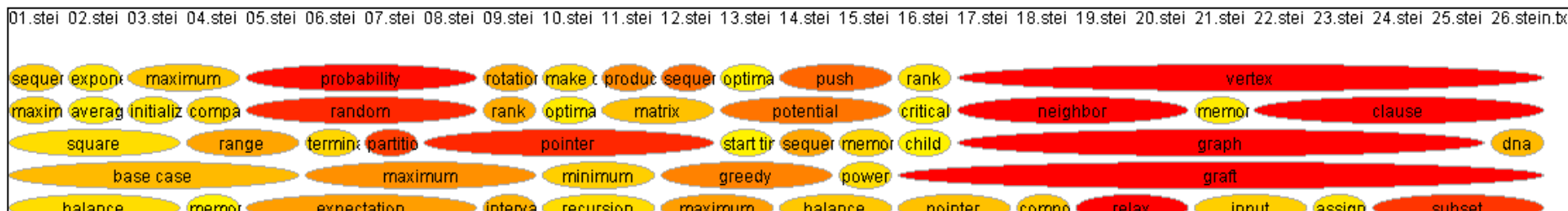
- Analysis of 273 transcripts (11 courses)
- Index Phrases
  - Unique per textbook indices: 253-4701
  - Unique per transcript: 8-98
  - Occurrence: 1 (35-50%), 5-50 (20-30%)
  - Unique per course: 40-347
- Word Pairs
  - $\approx$  10 times more than index phrases
  - Less meaningful for summaries

# *Interface: Parallels to a Camera*

- 3 visualization techniques
- Share 3 freely variable parameters:
  - Zoom: specificity of phrases
    - Occurrence of phrase across transcript
    - Range: Topic-specific to entirely thematic
  - Focus: emphasis of phrases
    - Range: 1 – N (lowest – highest occurrence)
  - Contrast: length of phrases
    - Range: 1 – K (K usually 3)

# Interface: Transcript Index Map

- Index phrases mapped to transcript
- Equivalent to textbook index
  - But: order by occurrence (highest near top)
  - Color coded (red→yellow = high→low occ.)
- Cross-reference terms among transcripts
  - Longer blobs for repeating phrases
- Greedy population of space near top



- Demo *Transcript Index Map*

# Interface: Chapter Transcript Match

- Transcripts mapped to textbook chapters
- Rows=transcripts, Columns=chapters
- Match score based on occurrences of terms between transcript and chapter
- Performance:

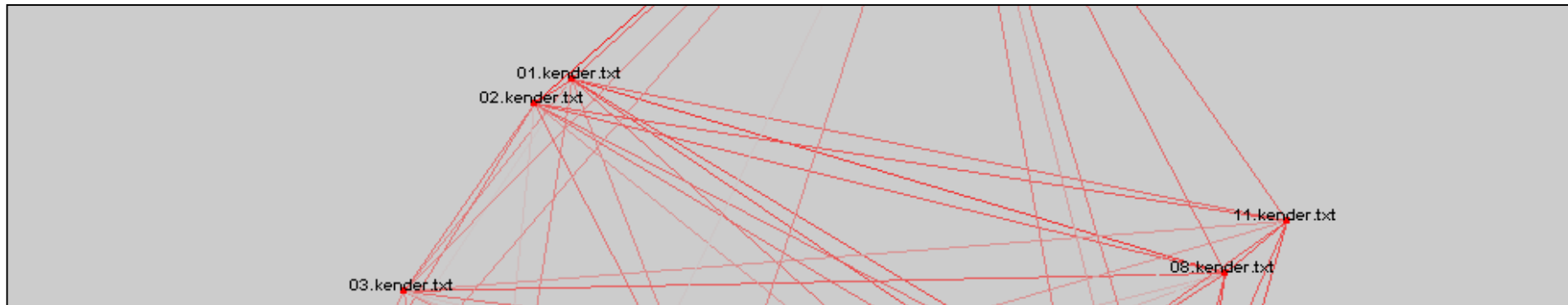
Data Set	Overall best accuracy over all Zoom levels
Index Phrases + Word Pairs	0.7
Word Pairs	0.66
Word Pairs from G <sup>2</sup>	0.63
Index Phrases	0.5

	ALL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
01.stein.txt	87...	48...	96...	49...	55...	47...	29...	46...	50...	52...	29...	64...	35...	35...	23...	12...	84...	53...	31...	17...	29...	44...	61...	41...	60...	46...	88...	41...	64...	61...	34...	85...	35...	71...	12...	69...	23...	24...	29...	42...
02.stein.txt	73...	24...	95...	37...	59...	28...	23...	62...	44...	40...	32...	43...	26...	31...	20...	62...	53...	41...	21...	27...	18...	27...	24...	18...	18...	22...	28...	34...	49...	60...	21...	49...	24...	54...	90...	57...	8...	25...	13...	27...
03.stein.txt	83...	32...	68...	15...	29...	21...	59...	32...	55...	31...	57...	46...	12...	11...	65...	74...	61...	62...	74...	54...	52...	42...	42...	34...	46...	15...	66...	43...	11...	38...	16...	57...	21...	43...	51...	33...	4...	63...	12...	57...
04.stein.txt	89...	42...	95...	22...	26...	35...	39...	64...	12...	80...	35...	67...	37...	34...	22...	84...	58...	39...	42...	27...	26...	38...	35...	20...	38...	22...	46...	48...	32...	46...	20...	67...	16...	64...	58...	40...	3...	33...	23...	31...
05.stein.txt	64...	18...	54...	8...	20...	96...	18...	65...	54...	59...	18...	48...	30...	21...	18...	48...	39...	27...	21...	15...	13...	15...	26...	14...	31...	15...	29...	17...	27...	35...	11...	49...	13...	34...	45...	40...	3...	24...	52...	22...
06.stein.txt	70...	14...	49...	14...	14...	10...	25...	45...	50...	32...	37...	66...	18...	39...	29...	73...	42...	30...	31...	21...	21...	44...	39...	24...	44...	13...	51...	17...	30...	49...	21...	60...	30...	59...	58...	48...	6...	24...	74...	42...
07.stein.txt	82...	11...	65...	23...	36...	76...	35...	12...	61...	77...	41...	61...	37...	38...	19...	66...	44...	38...	25...	22...	30...	35...	40...	23...	46...	14...	49...	19...	44...	63...	17...	59...	17...	56...	62...	45...	8...	39...	57...	34...
08.stein.txt	11...	24...	78...	18...	40...	69...	46...	46...	61...	37...	13...	10...	11...	10...	63...	10...	50...	76...	64...	57...	89...	91...	62...	25...	39...	22...	51...	20...	30...	42...	19...	48...	22...	63...	65...	54...	2...	44...	60...	17...

- Demo *Chapter Transcript Match*

# Interface: Transcript Similarity

- Cluster similar lectures (transcripts)
  - Dice distance of co-occurrence counts of selected index phrases
- Multidimensional scaling: |lectures|  $\rightarrow$  2D



- Demo *Transcript Similarity*



# Interface: Transcript Index Map

- Zoom=1; “highly topic-specific”

The screenshot displays the 'Transcript Analyzer' software interface. At the top, there are three control panels: 'Zoom' (Maximum # repeated Phrases across Transcript, slider at 1), 'Focus' (Minimum # repeated Phrases in Transcript, slider at 1), and 'Contrast' (Minimum # Words in Phrase, slider at 1). To the right is an 'Input Selection' panel with a list of files: transcriptIndexPhrase.txt, transcriptIndexPhraseDiff\_ABFreq.txt, transcriptIndexPhraseDiffA\_BFreq.txt, transcriptWordPair.txt, and transcriptWordPairG2.txt. Below these are two more sliders: 'Maximum gap between consecutive word appearances' (slider at 0) and 'Color 2: Maximum Frequency of phrase bubbles' (slider at 0). A 'Transcript Selection' panel on the far right lists files from 01.stein.txt to 26.stein.txt. The main area shows a grid of 26 columns (01.stei to 26.stein.txt) and several rows of words in colored bubbles. The words include: multipl, best case, stabilit, transitt, linearit, outer, list sea, rotator, make, simple, maxim, start tir, stack o, neighb, equiva, univers, succee, increm, relabel, comple, polyno, rejectic, hamilt, edge s, bottleneck, decom, factorial, main r, permut, depen, spine, interse, binary, fastest way, partial, huffma, equiva, sparse, relaxat, interm, literal, transpr, secon, inequality, recurre, recursion tree, decren, binomi, standard deviation, search, optimal, substructure, penalty, off line, union, arc, discret, identit, feasibi, equivalence, truth table, summation, merge, linearit, random number gener, query, consolidate, dictionary, aggregate analysis, task, polygo, auxiliary, quadratic, violation, mod, duality, red black tree, dense, dynamic table, tree edge, circuit, technology, red black, fractional, binary counter, adjacency matrix, fuzzy, inverse, representation, deletion, maximum value.

# Interface: Chapter Transcript Match

- Green=correct, Red=incorrect, Yellow=other possibility

Transcript Analyzer

**"Zoom"**

Maximum # repeated Phrases across Transcript

1 6 11 16 21 26 **10**

**"Focus"**

Minimum # repeated Phrases in Transcript

1 11 21 31 **1**

**"Contrast"**

Minimum # Words in Phrase

1 **1**

**Input Selection**

transcriptIndexPhrase.txt  
 transcriptIndexPhraseDiff\_ABFreq.txt  
 transcriptIndexPhraseDiffA\_BFreq.txt  
 transcriptWordPair.txt  
 transcriptWordPairG2.txt

Chapter Transcript Match

Transcript Index Map

Lecture Transcript Similarity

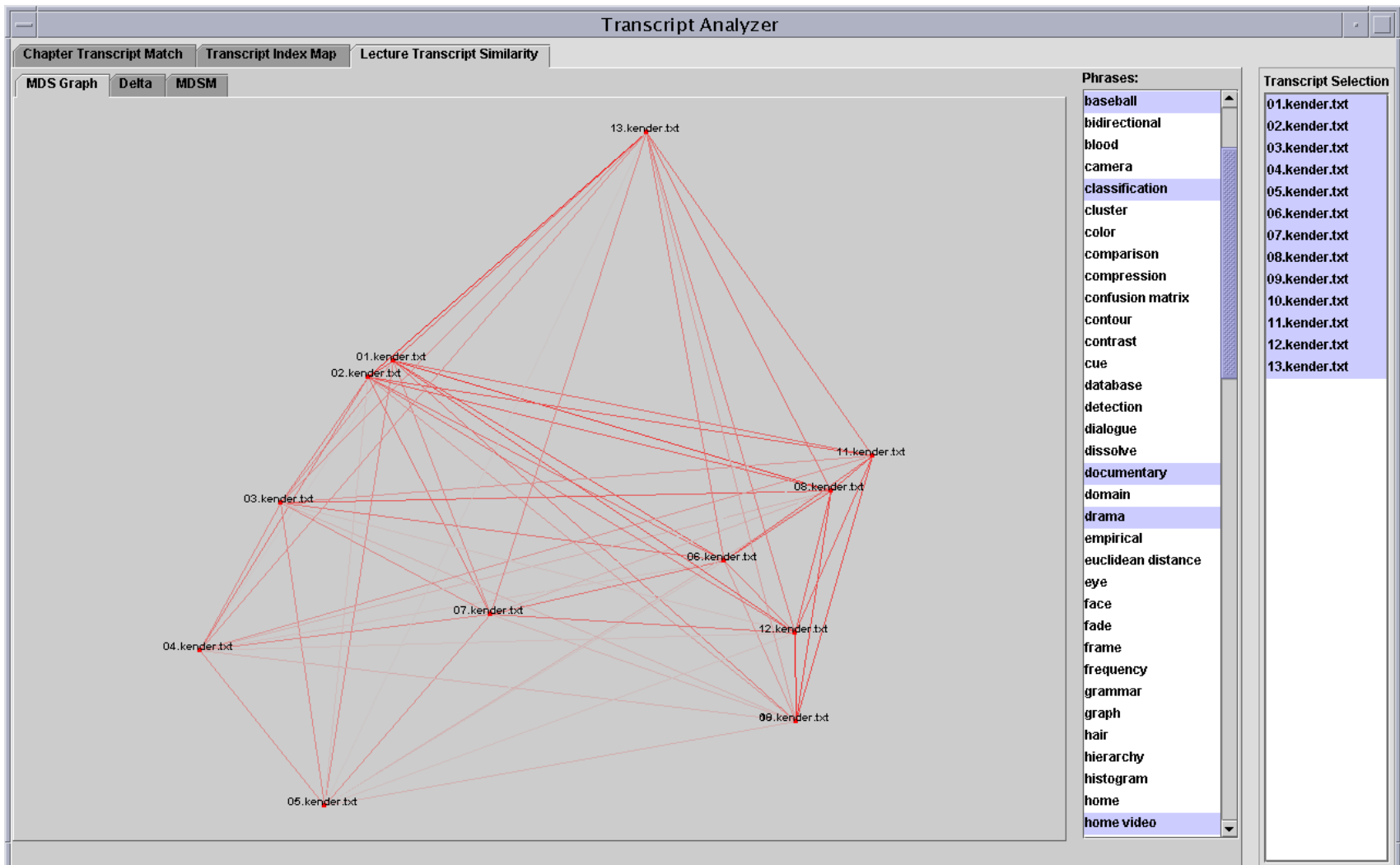
Transcript Selection

	ALL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
01.stein.txt	87...	48...	96...	49...	55...	47...	29...	45...	50...	52...	29...	64...	35...	35...	23...	12...	84...	53...	31...	17...	29...	44...	61...	41...	60...	45...	88...	41...	64...	61...	34...	85...	35...	71...	12...	69...	23...	24...	29...	42...
02.stein.txt	73...	24...	95...	37...	59...	28...	23...	62...	44...	40...	32...	43...	26...	31...	20...	62...	53...	41...	21...	27...	18...	27...	24...	18...	18...	22...	28...	34...	49...	60...	21...	49...	24...	54...	90...	57...	8...	25...	13...	27...
03.stein.txt	83...	32...	68...	15...	29...	21...	59...	32...	55...	31...	57...	46...	12...	11...	65...	74...	61...	62...	74...	54...	52...	42...	42...	34...	46...	15...	66...	43...	11...	38...	16...	57...	21...	43...	51...	33...	4...	63...	12...	57...
04.stein.txt	89...	42...	95...	22...	26...	35...	39...	64...	12...	80...	35...	67...	37...	34...	22...	84...	58...	39...	42...	27...	26...	38...	35...	20...	38...	22...	46...	48...	32...	46...	20...	67...	16...	64...	58...	40...	3...	33...	23...	31...
05.stein.txt	64...	18...	54...	8...	20...	96...	18...	65...	54...	59...	18...	48...	30...	21...	18...	48...	39...	27...	21...	15...	13...	15...	26...	14...	31...	15...	29...	17...	27...	35...	11...	49...	13...	34...	45...	40...	3...	24...	52...	22...
06.stein.txt	70...	14...	49...	14...	14...	10...	25...	45...	50...	32...	37...	66...	18...	39...	29...	73...	42...	30...	31...	21...	21...	44...	39...	24...	44...	13...	51...	17...	30...	49...	21...	60...	30...	59...	58...	48...	6...	24...	74...	42...
07.stein.txt	82...	11...	65...	23...	36...	76...	35...	12...	61...	77...	41...	61...	37...	38...	19...	66...	44...	38...	25...	22...	30...	35...	40...	23...	46...	14...	49...	19...	44...	63...	17...	59...	17...	56...	62...	45...	8...	39...	57...	34...
08.stein.txt	11...	24...	78...	18...	40...	69...	46...	46...	61...	37...	13...	10...	11...	10...	63...	10...	50...	76...	64...	57...	89...	91...	62...	25...	39...	22...	51...	20...	30...	42...	19...	48...	22...	63...	65...	54...	2...	44...	60...	17...
09.stein.txt	12...	19...	51...	15...	32...	22...	65...	39...	41...	27...	61...	42...	17...	28...	17...	13...	76...	78...	11...	67...	77...	81...	68...	49...	40...	25...	41...	10...	26...	48...	18...	50...	14...	55...	57...	40...	2...	11...	23...	82...
10.stein.txt	10...	53...	81...	21...	41...	37...	32...	37...	58...	45...	27...	34...	52...	56...	30...	25...	42...	24...	20...	32...	31...	58...	41...	79...	68...	65...	40...	32...	95...	29...	77...	38...	91...	13...	98...	1...	34...	21...	21...	
11.stein.txt	93...	26...	39...	5...	18...	17...	20...	29...	24...	25...	37...	40...	32...	32...	22...	23...	95...	28...	36...	28...	33...	35...	65...	21...	54...	51...	45...	16...	79...	99...	36...	80...	19...	88...	89...	64...	2...	30...	14...	25...
12.stein.txt	67...	22...	40...	11...	14...	24...	8...	15...	20...	26...	15...	36...	19...	24...	17...	10...	76...	25...	10...	15...	14...	20...	37...	31...	50...	24...	38...	22...	31...	66...	13...	42...	33...	44...	71...	57...	3...	22...	27...	15...
13.stein.txt	77...	18...	44...	10...	23...	25...	26...	24...	43...	23...	41...	27...	45...	50...	31...	11...	16...	47...	24...	32...	34...	27...	64...	45...	50...	25...	27...	14...	20...	63...	21...	52...	30...	50...	66...	66...	2...	35...	17...	35...
14.stein.txt	76...	8...	33...	1...	27...	21...	22...	28...	30...	15...	84...	32...	55...	69...	41...	85...	74...	17...	53...	46...	63...	85...	33...	27...	27...	9...	49...	22...	15...	28...	17...	36...	15...	48...	32...	32...	2...	41...	23...	51...
15.stein.txt	71...	17...	68...	30...	29...	20...	24...	32...	31...	20...	26...	71...	32...	36...	29...	82...	26...	16...	34...	22...	37...	51...	24...	12...	42...	13...	55...	17...	16...	42...	15...	52...	25...	52...	53...	34...	18...	26...	30...	46...
16.stein.txt	10...	24...	45...	15...	29...	31...	52...	37...	31...	28...	10...	72...	83...	88...	72...	10...	90...	80...	60...	81...	82...	13...	60...	49...	52...	20...	72...	36...	30...	50...	24...	71...	31...	11...	83...	55...	4...	81...	22...	63...
17.stein.txt	11...	19...	50...	20...	28...	27...	42...	47...	43...	34...	49...	51...	62...	81...	56...	10...	79...	73...	47...	42...	60...	75...	30...	11...	13...	79...	14...	24...	40...	43...	33...	48...	22...	67...	13...	87...	7...	11...	17...	61...
18.stein.txt	94...	10...	40...	5...	13...	25...	9...	26...	28...	21...	21...	24...	38...	38...	34...	66...	97...	31...	30...	18...	27...	37...	23...	14...	10...	45...	10...	21...	16...	28...	12...	27...	18...	70...	82...	67...	0...	74...	8...	21...
19.stein.txt	10...	19...	46...	8...	21...	28...	53...	23...	38...	22...	55...	50...	67...	71...	66...	97...	83...	67...	43...	45...	69...	82...	11...	12...	11...	68...	11...	9...	35...	60...	16...	40...	19...	54...	85...	83...	2...	68...	15...	46...
20.stein.txt	72...	21...	39...	10...	12...	21...	34...	21...	19...	15...	35...	23...	20...	40...	28...	69...	54...	34...	25...	26...	30...	42...	97...	59...	12...	70...	99...	12...	11...	32...	9...	22...	15...	68...	78...	43...	2...	45...	17...	29...
21.stein.txt	96...	44...	79...	16...	33...	31...	34...	42...	60...	35...	47...	42...	53...	53...	20...	15...	73...	39...	40...	24...	18...	37...	93...	50...	95...	83...	92...	31...	61...	70...	29...	69...	23...	85...	17...	85...	6...	50...	23...	35...
22.stein.txt	73...	25...	43...	11...	27...	16...	17...	21...	22...	18...	20...	24...	24...	55...	12...	78...	55...	18...	16...	19...	18...	28...	32...	37...	67...	34...	58...	16...	21...	67...	22...	31...	32...	37...	19...	84...	3...	50...	16...	26...
23.stein.txt	63...	10...	20...	5...	11...	15...	10...	8...	11...	9...	7...	31...	10...	27...	7...	48...	57...	14...	5...	7...	11...	19...	67...	55...	60...	35...	66...	3...	8...	37...	5...	35...	9...	46...	18...	10...	0...	57...	8...	13...
24.stein.txt	10...	26...	38...	22...	39...	25...	14...	31...	42...	24...	30...	34...	39...	31...	24...	13...	11...	30...	29...	32...	10...	61...	87...	54...	11...	14...	62...	98...	21...	56...	20...	62...	21...	18...	9...	81...	30...	11...		
25.stein.txt	90...	23...	36...	18...	22...	43...	26...	28...	36...	40...	33...	42...	39...	41...	25...	89...	93...	30...	29...	35...	31...	48...	75...	72...	83...	44...	83...	9...	28...	89...	9...	52...	16...	61...	15...	19...	8...	60...	44...	24...
26.stein.txt	94...	13...	69...	15...	39...	21...	31...	40...	49...	41...	58...	47...	19...	31...	31...	97...	90...	64...	24...	30...	30...	63...	57...	47...	53...	38...	70...	22...	38...	39...	24...	52...	12...	71...	94...	16...	13...	52...	26...	51...

Resize Columns to fit

# Interface: Transcript Similarity

- Course on “Visual Databases”



# *Future Investigations*

- Incorporate tools into previously developed lecture browser
- User studies: usability of interface, ease of finding information
- Extend interface: provide index phrase layout on lecture level
- Extend indexing and visualization from intra-lecture to intra-course

*Thank you!*