Textbook Summary Spring 2023

Dennis, Christina

Goals and Motivation

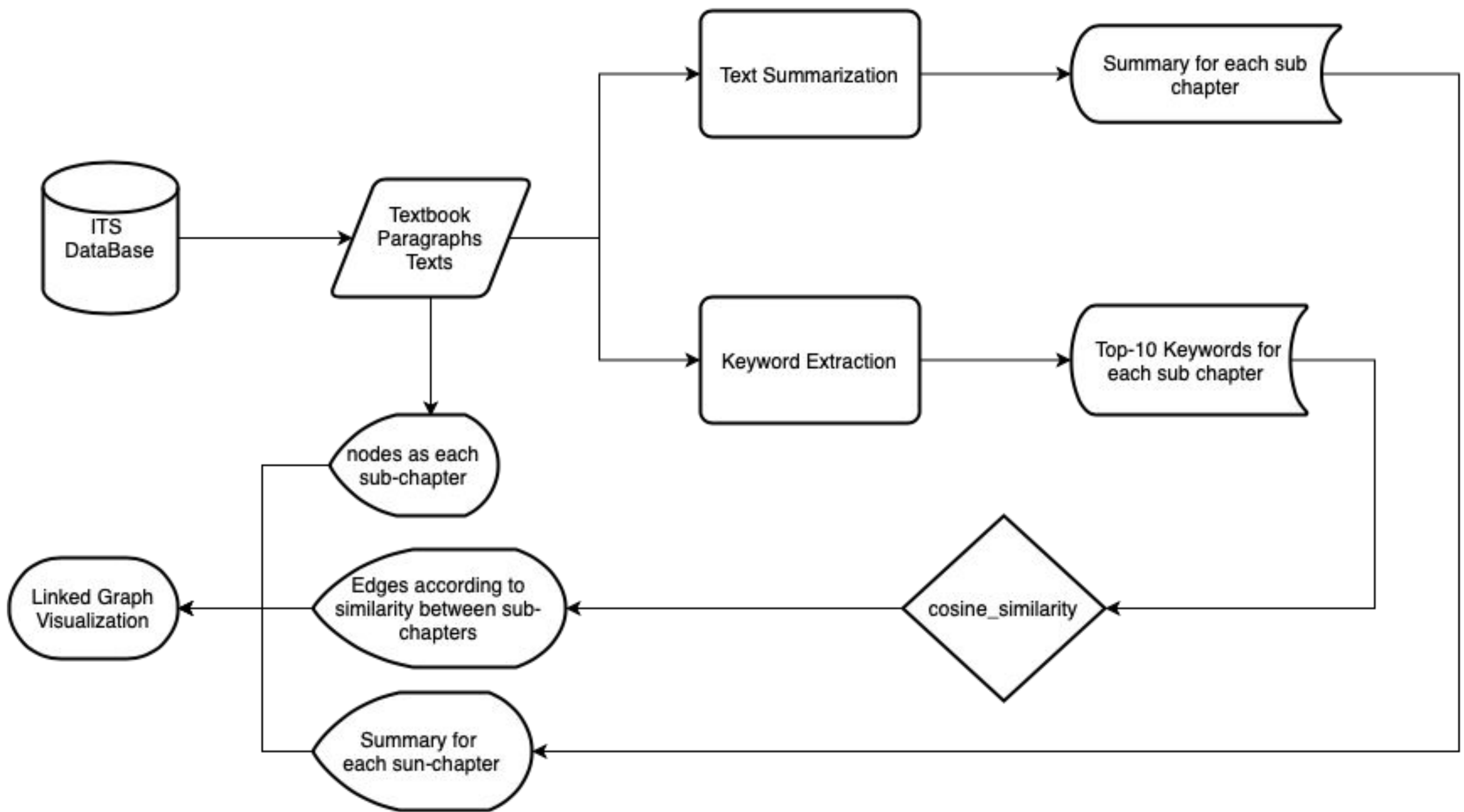
1. Student-centric textbook navigation/quality of life.
    a. Study help
    b. Deeper understanding/inquisitiveness

2. Creating/Visualizing a knowledge structure/hierarchy of a textbook.
    a. Representing an entire textbook with one screen
    b. Interconnectedness of a subject expertise

Approach

- Create textbook summaries for each sub-chapter (1.1, 3.10, …)
- For each sub-chapter, extract the keywords
- Use these keywords to calculate the similarity between each pair of sub-chapters
- Visualize similar sub-chapters using a vertex-edge graph model
- Allow users to view the summaries for each sub-chapter.

Flow Chart for Textbook Summary Visualization



Textbook Summarization

1. Data processing: convert paragraphs to sentences
2. Create Frequency Table for words without stop words
3. Calculate sentence scores
4. Use top-n scores to figure the summary sentences

Ch. 4 summary


```
 '4.0': '  CHAPTER 4 Sampling and Aliasing This chapter is concerned with the conversion of signals between the analog (continuous-time) and digital
(discrete-time) domains. The primary objective of our presentation is an understanding of the Sampling Theorem, which states that when the sampling
rate is greater than twice the highest frequency contained in the spectrum of the analog signal, the original signal can be reconstructed exactly
from the samples. The process of converting from digital back to analog is called reconstruction. The reconstruction process is basically one of
interpolation because we must “fill in” the missing signal values between the sample times by constructing a smooth curve through the discrete-time
sample values.',
 '4.1': ' It is also common to refer to such signals as analog signals because both the signal amplitude and the time variable are assumed to be
real (not discrete) numbers. It is a normalized version of the continuous-time radian frequency with respect to the sampling frequency. It is an
alias. The frequency of is, while the frequency of is. There are many more frequency aliases as the following problem suggests. Show that is an
alias of. There are other aliases. Show that the signal is an alias of the signal.',
 '4.2': ' These are what we have called the principal alias frequencies, and in Fig.~※ these are the frequency components that fall inside the
dashed box shown in the bottom panel, i.e.,. These are at, so we calculate the output spectrum lines at Hz. The analog frequency of the
reconstructed output will be Hz. The output frequency will be equal to the input frequency. Now consider the input frequency increasing from to.',
 '4.3': ' Six successive positions of the spot for a very high flashing rate. This is aliasing because the spot makes exactly one revolution between
flashes and therefore is always at the same position when illuminated by the strobe. This is not the only flash rate for which the spot will appear
to stand still. Six successive positions of the spot for a flashing rate that aliases the spot motion. The frequency of the motor rotation is
constant, as is the radius. The frequency is the rotation rate of the motor in rpm. What is the maximum flashing rate for this case?',
 '4.4': ' Interpolation with Pulses How does the D-to-C converter work? Four different pulses for D-to-C conversion. D-to-C conversion using a
square pulse. Thus, using a square pulse in is D-to-C conversion, but not ideal D-to-C conversion. D-to-C conversion using a triangular pulse, or
linear interpolation. D-to-C conversion using a cubic-spline pulse. D-to-C conversion using a square pulse. D-to-C conversion using a cubic-spline
pulse at samples/sec.',
 '4.5': '  The Sampling Theorem This chapter has discussed the issues that arise in sampling continuous-time signals. Sampling and reconstruction
system. Such signals could be represented as where each of the individual signals is of the form and it is assumed that the frequencies are ordered
so that and. If we sample the signal represented by and, we obtain where with.',
 '4.6': '  Summary and Links This chapter introduced the concept of sampling and the companion operation of reconstruction. With sampling, the
possibility of aliasing always exists, and we have created the strobe demo to illustrate that concept in a direct and intuitive way. Synthetic
strobe demos produced as MATLAB movies. Reconstruction movies that show the interpolation process for different pulse shapes and different sampling
rates. A sampling and aliasing MATLAB GUI called con2dis, shown in Fig.~※. Aliasing and folding movies that show how the spectrum changes for
different sampling rates.',
```

Keyword each paragraphs

1. Vectorizer text to numbers
2. Get important words as unigram keyword candidates
3. Encode doc and candidates in Sentence Transformer
4. Find 20 candidates most similar to the document
5. Pick up a combination of 10 least similar keywords

Keywords for Ch. 4

4.0	['reconstruction', 'count', 'arithmetic', 'engineering', 'compilers', 'polynomials', 'sampled', 'hardware', 'cd', 'audio']
4.1	['pipelining', 'microprocessor', 'microprocessors', 'audio', 'polynomials', 'sampled', 'sinusoidal', 'shading', 'arithmetic', 'mathematical']
4.2	['sinusoidal', 'sinusoids', 'compilers', 'cd', 'integer', 'computing', 'filtering', 'multiplications', 'audio', 'polynomials']
4.3	['microprocessors', 'easily', 'sinusoidal', 'computer', 'mathematical', 'shading', 'pipelining', 'rich', 'fast', 'tools']
4.4	['polynomial', 'linear', 'invariance', 'optimized', 'converter', 'reconsider', 'polynomials', 'algorithms', 'sampled', 'audio']
4.5	['mathematically', 'optimized', 'mathematical', 'filtering', 'reconsider', 'polynomials', 'insight', 'sampled', 'audio', 'algorithms']
4.6	['classes', 'sampled', 'tools', 'hardware', 'mathematical', 'aliasing', 'lab', 'homework', 'demo', 'microprocessors']

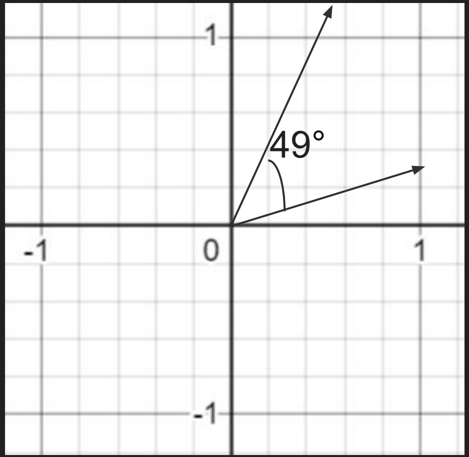
Sub-Chapter Similarity

1.4
Sinusoidal 1.0
Signal 0.50
Mathematical 0.30

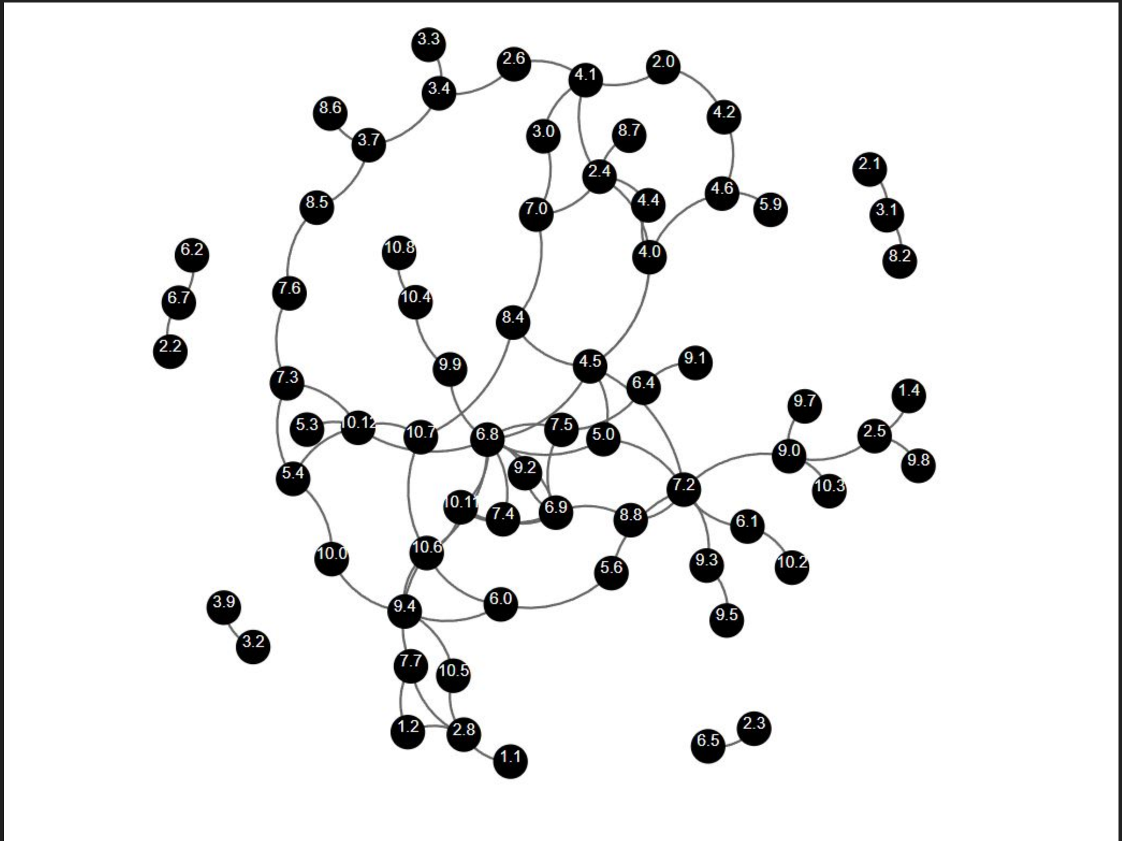
1.8
Signal 1.0
Frequency 0.80
Sinusoidal 0.25

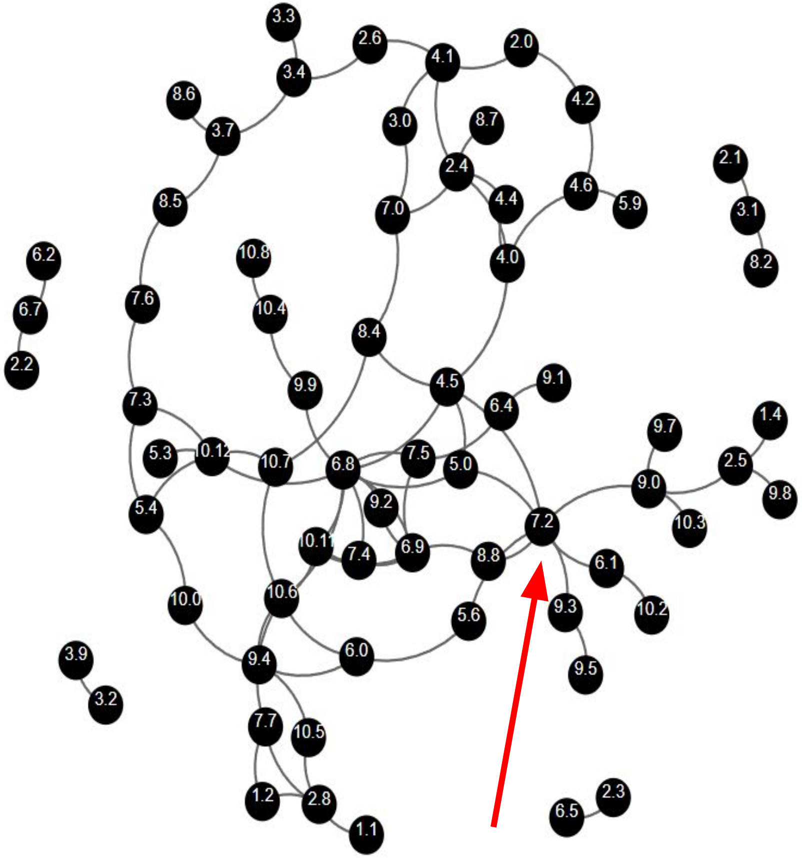
$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|}$$

<0.5, 1.0>, <1.0, 0.25> -> similarity = 0.651

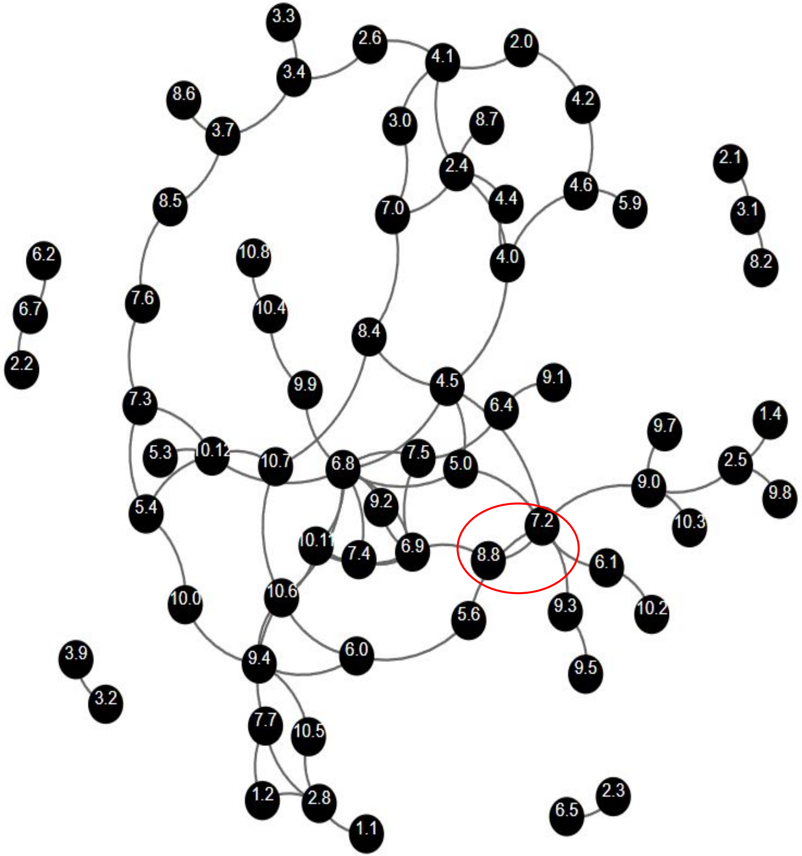


Visualization

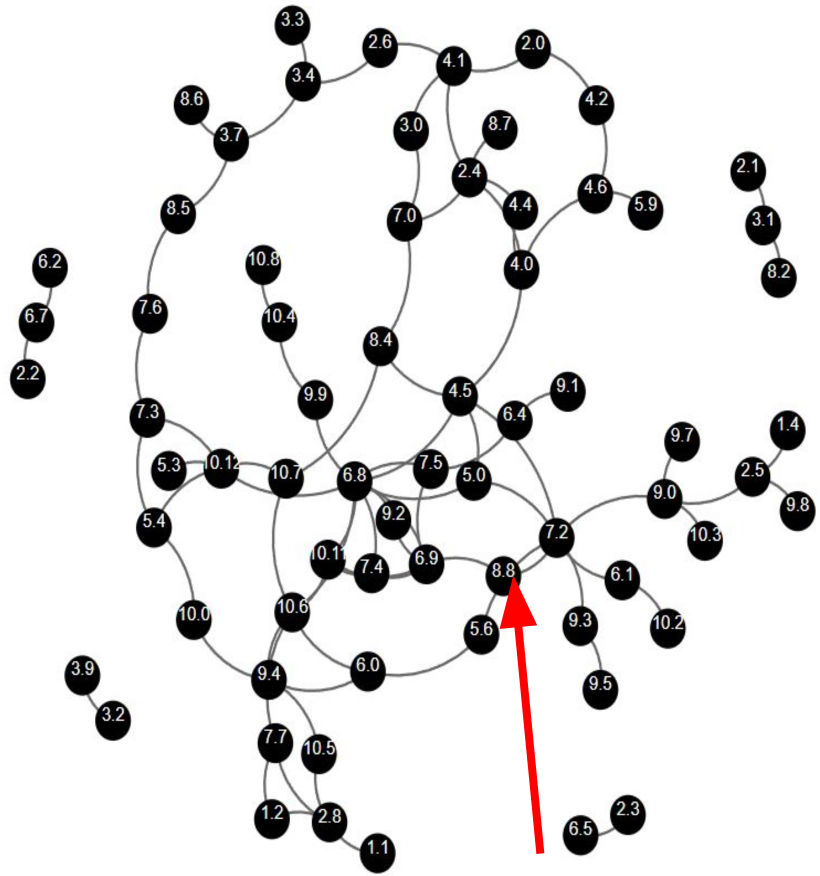




EXAMPLE: DTFT of an FIR Filter The following FIR filter has a finite-length impulse response signal: Each impulse in is transformed using, and then combined according to the linearity property of the DTFT which gives Uniqueness of the DTFT The DTFT is a unique relationship between and ; in other words, two different signals cannot have the same DTFT. EXAMPLE: DTFT of Complex Exponential? The result of the inverse DTFT is the discrete-time signal where. Riemann sum approximation to the integral of the inverse DTFT. Here, which is the DTFT of.



EXAMPLE: DTFT of an FIR Filter The following FIR filter has a finite-length impulse response signal: Each impulse in is transformed using, and then combined according to the linearity property of the DTFT which gives Uniqueness of the DTFT The DTFT is a unique relationship between and ; in other words, two different signals cannot have the same DTFT. EXAMPLE: DTFT of Complex Exponential? The result of the inverse DTFT is the discrete-time signal where. Riemann sum approximation to the integral of the inverse DTFT. Here, which is the DTFT of.



The window is zero for and. DTFT of Windows Windows are used often in spectrum analysis because they have desirable properties in the frequency domain. Figure~※(a) shows plots of the magnitude of the DTFT for two Hann filters with lengths and. DTFT of two Hann windows; length (colored lines), and length (black/gray lines) (a) as computed by 1024-point FFT with zero padding and (b) as plotted with DFT values reordered to place in the middle. Figure~※(b) shows that the DTFT of the Hann window is concentrated near. DTFT of windowed sinusoid via 1000-pt. DFT for (a) rectangular window (peak heights are 20), and (b) Hann window (peak heights are 10.25). DFT for (a) Hann window scaled by, and (b) Hann window scaled by.

Conclusion/Possible Future Improvements

- Adjacency matrix to better visualize connections
- More relevant keywords
  - Stemming
  - Comparisons of NLP approaches
  - Index (Too many filler words)
- Filtering based on chapter
- Similarity between keywords