

Mohamed Elhoseiny and Ahmed Elgammal  
Department of Computer Science, Rutgers University

## Introduction

With today's vast and rapid increase in the use of electronic document readers, smart phone, tablets, there is an eminent need to develop tools for visualizing and summarizing textual contents. MindMapping [2] is not only a note-taking tool, but also a very powerful tool for text summarization and visualization. Converting a text paragraph to a MindMap would provide an easier way to visually access the knowledge and ideas in the text.

**William Shakespeare (1564-1616)**

Very little is officially known about Shakespeare, but scholars have pieced together a reasonably comprehensive picture of his life from his marriage to Anne Hathaway in 1582 to his Christenings of his 3 children. And most reputable critics agree to him the authorship of the major portion of thirty eight of the world's most respected plays several excellent poems, and some 154 sonnets.

William achieved success largely on his own. He apparently never attended college. Successive purchases and sales of agricultural products and parcels of land near Stratford must have provided Shakespeare with greatly increased capital, which, when reinvested paid him steadily income for many years. This gave him the freedom and time to concentrate on his first love: acting and writing.

In 1594-1595 William performed before Queen Elizabeth and his name became widely recognized. Shakespeare grew in public stature when he became one of the owners of London's Globe Theatre in 1599.

For Study, Shakespeare's works can be divided into six separate, somewhat chronological sections. Early works, Major Histories, The problem plays, Tragedies, The Roman Plays, and The Late Romances. His comedies seem to be interspersed throughout these divisions.

## Applications

- 1) Audio of a lecture to Text to MindMap.
- 2) Automatic Generation of MindMap Presentations.
- 3) A novel way to visualize and summarize text documents for electronic book readers.

### Areas

- ❖ Education
- ❖ Presentation
- ❖ Electronic document readers
- ❖ Human computer interaction



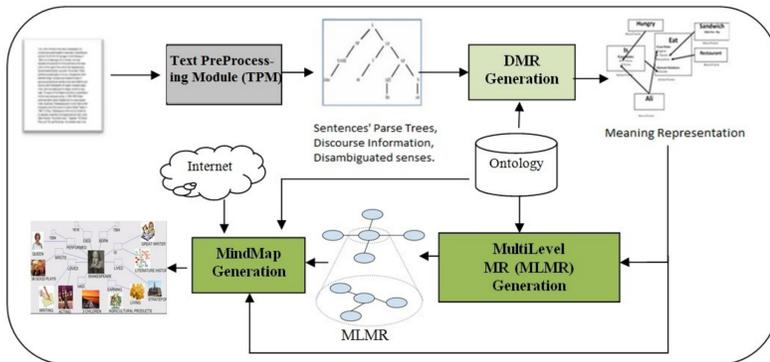
## Why Hierarchical Mind Maps?

Single level MindMap defies the purpose of comprehension speed, simplicity, and clarity for larger text.



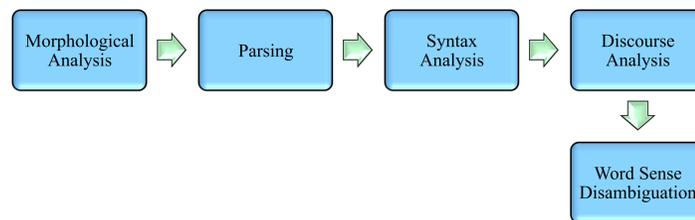
## System Architecture

This figure illustrates the architecture of English2MindMap system, where the blocks are color-coded according to the contribution in each block. Gray blocks almost used existing approaches and/or technologies (e.g., the TPM), while green blocks constitute the main contribution. The following subsections describe the functionality of each component.

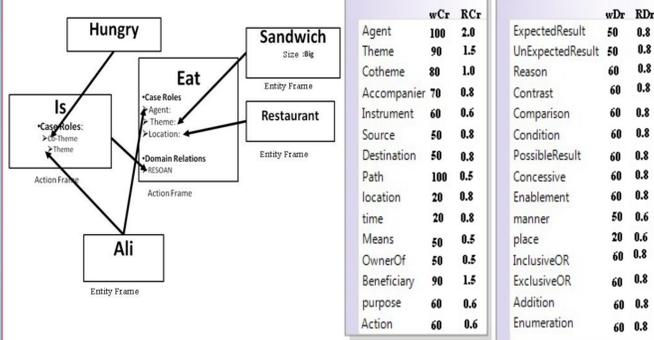


## Text Preprocessing Module (TPM)

This is the first phase in the system, it extracts sentence-wise information from raw text. It takes as an input the English plain text, and then generates for each sentence a parse tree, discourse analysis results, and the intended sense for each word in the sentence. In this stage, we utilized the current approaches for such a well-studied phase in NLP.



## DMR Generation



## MLMR Generation

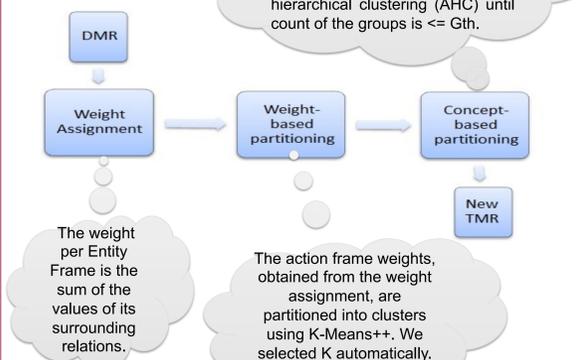
### ALGORITHM 1: MLMR Generation(DMR)

```

topMR ← MRSA(DMR)
Display MindMap of topMR to the user and Highlight group frames
while true do
  Wait until user select group frame, exit if user choses to exit
  currentMR ← MRSA(regionInDMR(Selected Frame))
  Display MindMap of currentMR to the user and Highlight group Frames
end while
  
```

### MRSA Algorithm

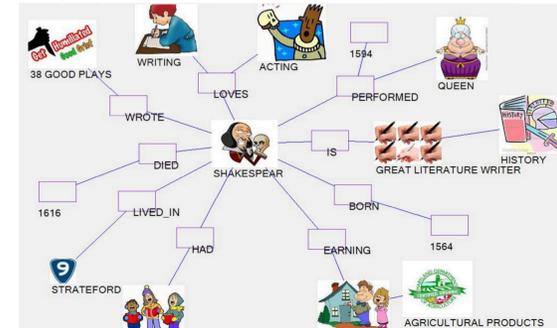
- 1) Group frames of the exact concept (i.e. ontological distance = 0).
- 2) If the count of the groups in 1 is > Gth (We used Gth=3), perform agglomerative hierarchical clustering (AHC) until count of the groups is <= Gth.



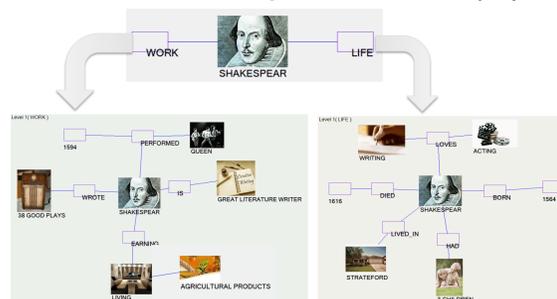
## MindMap Generation

This phase converts either the DMR or MLMR to a MindMap that contains images for visual frames then it automatically allocate it on the screen

### Singe Level MindMap from Pure Text (ClipArt)



### Hierarchical MindMap from Pure Text (All)



## Experiments (4900 Mturk Responses)

- (1) Correct?, (Regardless the pictures) ? Grade.
- (2) Relevant Pictures ? Grade.
- (3) How many missing Actions in the MMap)?.
- (4) How many missing Entities?.
- (5a) How Many redundant frames ?
- (5b) Satisfactory Hierarchy? Grade.

| Experiment 1 |      |      |      |      |      |
|--------------|------|------|------|------|------|
|              | Q1   | Q2   | Q3   | Q4   | Q5   |
| Mean         | 4.42 | 4.34 | 0.54 | 0.59 | 0.48 |
| StdDev       | 0.84 | 0.87 | 1.58 | 1.74 | 1.68 |
| SR           | 0.86 | 0.84 |      |      |      |

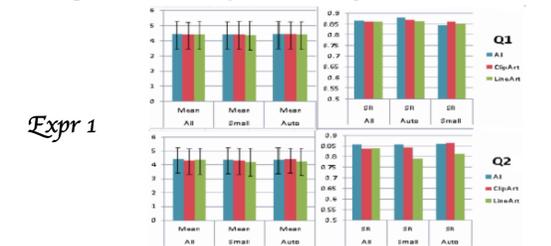
  

| Experiment 2 |      |      |      |      |      |
|--------------|------|------|------|------|------|
|              | Q1   | Q2   | Q3   | Q4   | Q5   |
| Mean         | 4.51 | 4.27 | 0.45 | 0.5  | 0.23 |
| StdDev       | 0.65 | 0.82 | 1.25 | 1.47 | 0.58 |
| SR           | 0.92 | 0.79 |      |      |      |

| Experiment 3 |      |      |      |      |      |
|--------------|------|------|------|------|------|
|              | Q1   | Q2   | Q3   | Q4   | Q5   |
| Mean         | 4.48 | 4.34 | 0.59 | 0.61 | 0.25 |
| StdDev       | 0.87 | 0.94 | 1.51 | 1.55 | 0.93 |
| SR           | 0.89 | 0.84 |      |      |      |

### Image Search Specs analysis



Expr 1

Expr 2

|         | Mean | StdDev | SR    | Mean | StdDev | SR   |
|---------|------|--------|-------|------|--------|------|
| All     | 4.56 | 0.63   | 0.937 | 4.31 | 0.825  | 0.8  |
| LineArt | 4.53 | 0.61   | 0.939 | 4.24 | 0.827  | 0.76 |
| ClipArt | 4.46 | 0.71   | 0.91  | 4.26 | 0.825  | 0.81 |

## Conclusions and Future Work

We have designed and implemented an automated tool that takes English text as input and generates a Mind Map visualization out of it. The system was comprehensively tested under different parameter settings by MTurk Human Subjects and high satisfaction rates have been recorded. Hence, we aim to extend the system such that it's reliable in handling very large text (e.g., a book) and also to try different approaches of Concept Combination. We will also work on enhancing the performance of the system to handle large text in reasonable time (i.e. ML MindMap Generation of a 250 word document take about 40 seconds on a 3.3GHZ-2GB machine).

### References

- [1], M. Elhoseiny, A. Elgammal, "English2MindMap: an Automated System for MindMap Generation from English Text," ISM, Dec, 2012.
- [2] B. B. Tony Buzan and J. Harrison, The Mind Map Book. BBC Archive, 2010.