CHAPTER VII

CONCLUSIONS AND FUTURE RESEARCH

7.1 CONTRIBUTIONS, STRATEGIES FOR EVOLVING CONCEPT EXTRACTION TECHNIQUE

Major Contribution:
The thesis has evolved a technique for extracting concepts of ‘Probability Theories’ through automatic computational methods, which can be adopted in e-learning environments. The major objective of this technique is to maintain efficient extraction means that includes reducing the computational resources like the CPU, while retaining the reliability in correctly extracting user specified concepts.

This technique has been validated under three distinctly different e-user levels, namely ECE, Bio-Tech and MBA for extracting concepts of ‘Probability Theories’ for three distinctively different situations and requirements.

7.1.1 Major Findings from survey

1. The current technical professional students are capable of abstractly distinguishing domain concept words from that of rhetoric concept words, and thereby accept ‘rhetoric portrayal abilities’ as ‘learning or comprehending concepts’.
2. Most of the technical professional students want numerical problem examples/assignments to be representative of their own discipline problems with appropriate mathematical treatments.

7.1.2 Observations, Findings and Inferences drawn from the social studies

1. More than 50% of the respondents have expressed that their current mathematical contents are not relevant to their professional disciplines.
2. Majority of the respondents are in favour of the future mathematical contents should be relevant to their professional disciplines.
3. Majority of the respondents have expressed that the mathematics contents must include conceptual description in addition to mere problem solving examples.
4. The existing instructional approaches of mathematical contents are not clear and not modular.
5. Most of the respondents have expressed that the mathematical instructional contents need to be highly structured.

6. On the identification of domain and non domain words, the respondents have observed that the current instructional materials have mixed words.

7. It is inferred that the current professional students are capable of segregating rhetoric words from domain words from their mathematical contents.

7.2 CONTENT ANALYTICAL STUDIES

7.2.1 Major conclusions from the analytical study

1. Merrill’s portrayal based approach has proved to be yielding concept words in terms of rhetoric portrayals that are expressed through concept words for extracting rhetoric concepts.

2. Content analytical studies have clearly demonstrated the classification of rhetoric concepts through quantitative classifications.

7.2.2 Observations, Findings and Inferences from the content studies

1. Both the content classification techniques namely the Bloom’s taxonomy as well as portrayal based classifications have quantitatively different rhetoric concepts of domain dependent documents.

2. It is observed that the rhetoric concept words repeated in several concept classifications of Bloom, while it is found to be mostly non-repetitive in portrayal based classifications.

7.3 EXPERIMENTAL STUDIES ON CONDITIONAL PROBABILITIES

Apart from independent probability analysis, mutually inclusive and exclusive conditional probability criteria for concept extraction were planned according to one of the research objectives. The mutual inclusive probability was demonstrated from Bloom’s hierarchical model and the mutually exclusive probability from the researcher’s three derived portrayals has been experimented with conclusions drawn from such arrangements are presented below.
7.3.1 Major conclusions from the experimental study

1. The introduction of Pedestal Concept-structure Document (PCD), a novelty introduced by the research has proved to be most advantageous for successful extractions of concepts, when the concepts have been delimited to portrayal based ones.

2. Most of the non domain confounded words were found to be rhetoric concept in nature, while the linking words were semi dependent or action oriented words on the domain dependent words.

3. The confounded words of documents (namely the PCDs) have proved to be effective in representing portrayal concepts, which were helpful in concept extractions (a hypothesis assumed by the research has been validated).

4. The repetitive action verbs of Bloom’s taxonomies proved to be ineffective for mutually exclusive probability condition, whereas the non repetitive portrayal words of the researcher’s derived portrayals have been found to be effective for mutually exclusive conditional probability.

5. The thesis has clearly demonstrated the effectiveness of Naive Baye’s conditional theory when applied in exclusive probability conditions.

6. Portrayal based classifications proved to be effective in concept extractions (in terms of accuracy and acceptance levels).

7. Mutually exclusive conditional probability approach is the most suitable concept categorization.

7.3.2 Observations, Findings and Inferences from the experimental studies on concept extractions

1. It is inferred from the experiments that the portrayal based concept extractions are closer to the actual concepts than actual and predicted by Bloom’s concepts.

2. In most of the case studies, it is found that the probability values increased when link words are removed from the documents.

3. The mutually inclusive conditional probability values of Bloom’s taxonomy words consistently performed lower than mutually exclusive probability of portrayal concepts.
4. The percentage of deviations in mutually inclusive conditional probability of Bloom’s concepts was found to be very high when compared with the deviations of portrayal based mutually exclusive conditional probability.

5. Significant difference found between actual Bloom’s concepts presented in documents from that predicted by the mutually inclusive conditional probability approach.

7.4 PARAMETRIC STUDIES ON PCD EXTRACTIONS

Parametric study was aimed at determining best possible storage conditions of PCDs, for efficient extraction of PCDs. Four types of conditions (chapter VI) have been considered. Three disciplines namely B.Tech (ECE), B.Tech (Bio Tech) and MBA (PG Management) programmes that have Probability and Statistics as subjects (courses) in their studies have been considered for the study. The parameters for determining efficiency through the experiments are CPU consumed by the system while searching for the appropriate PCDs (not to get confused with concept searching); Probability values and percentage of successful retrieval as per user attempt. Eleven episodes (chapter VI) have been delimited for the study. As explained in the last chapter, several small PCDs have been developed for the experiments. Conclusions that have been derived out of these detailed experiments are brought out below.

7.4.1 Major conclusions from the parametric study

1. The elaborate social study and the experiments confirmed the advantage of arranging portrayal wise PCDs for successful extractions for not only concept extraction but also the extractions of PCDs.

2. From the overall study it is concluded that the best condition for efficient extraction of PCDs is the fourth condition, namely single storage with indexed PCDs with discipline, and portrayal type. Even though the CPU time consumed by the system was not the least, the successful extractions were almost 100%.

3. It is found from the detailed parametric study experiments, that the computational (CPU) time consumed by the system is directly proportional to the number of (small) PCDs of a
particular condition; say a discipline or a particular portrayal. Hence it is concluded that
reduction of CPU time alone cannot contribute to the efficiency of PCD extractions.

4. No significant difference is found between independent probability and conditional
probability values in the case of PCD extractions, unlike concept extractions.

7.4.2 Observations, Findings and Inferences from the Social Study

1. Majority of the ECE respondents prefer their own discipline wise real world problems in
treating probability and statistics mathematical courses.
2. Preference for ‘Explicit’ portrayal was more by the ECE respondents.
3. Preference for the rest of the portrayals was not much by the ECE respondents.
4. The majority of the Bio Tech. respondents’ preferred for discipline wise real world
problems more than ECE respondents.
5. There is no clear cut preference for any particular portrayal by the Bio-Tech respondents.
6. The preference for discipline wise real world problem by the MBA respondents was
significantly higher than other discipline respondents.
7. The preference for ‘Explicit’ portrayal by the MBA respondents was significantly less
than other discipline respondents.
8. The preference for ‘Analytical’ portrayal was significantly higher for MBAs than other
discipline respondents.

7.4.3 Observations, Findings and Inferences drawn from Experimental Studies

1. It is observed from the experiments that the CPU time consumption is greatly influenced
by the number of PCDs stored in a database and not caused either by portrayal class by
episodic contents.
2. The second largest successful and acceptable PCD extraction is ‘Condition 3’ namely
discipline wise grouping of PCDs.
3. There is no significant difference between success rates of small or larger number of
trials. The percentiles are more or less the same for both these extremes.
4. The successful extraction rates are significant to the discipline and portrayal preference.
For example the portrayal ‘Analytical’ was preferred mostly by MBA (Management)
respondents while the portrayal ‘Informative’ was largely preferred by the B.Tech (ECE)
respondents.
5. There is a significant difference between the probability values of ‘Condition 4’ and the rest of the conditions.
6. There is a significant difference between the largest CPU consumption condition (namely 3: discipline wise grouping) of ECE discipline and the rest of the disciplines.
7. As a general thumb rule the CPU consumption will be the largest in indexed based PCDs.
8. No specific advantage is noticed by arranging the PCDs according to episodes (condition 2).

7.5 UTILITY VALUES OF RESEARCH AND RECOMMENDATIONS
The proposed technique can be used as an automated procedure for extracting concepts of ‘Probability Theories’ for different learner levels (disciplines) efficiently and reliably.
The research findings, particularly from the surveys are indicative of different user requirements in dealing with the subject: “Probability Theories”, specific to their disciplines.

7.5.1 SUGGESTIONS FOR FUTURE RESEARCH
This research work can be extended to non similar mathematical subjects like “Series and Functions” that are applied to different disciplines so as to compare with these results. Reasons for any deviations could be debated.
The combination of e-contents in the form of modules could be automated through mathematical formulae such as ‘series’ so as to make the automation more efficient.