Knowledge environment for personalizing e-learning

Teo, Chao Boon

2007

Teo, C. B. (2007). Knowledge environment for personalizing e-learning. Doctoral thesis,
Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/3544>

<https://doi.org/10.32657/10356/3544>

Nanyang Technological University

Downloaded on 26 Apr 2023 20:47:18 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

A KNOWLEDGE ENVIRONMENT FOR
PERSONALIZING E-LEARNING

**A KNOWLEDGE ENVIRONMENT FOR
PERSONALIZING E-LEARNING**

TEO CHAO BOON

TEO CHAO BOON

SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

2007

2007

A Knowledge Environment for Personalizing E-learning

Teo Chao Boon

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

2007

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

07 December 2007

Date

Teo Chao Boon

Acknowledgements

I would like to give my sincere and deepest thanks to my supervisors, Professor Robert Gay Kheng Leng, and Professor Habib Mir M. Hosseini, for their patient guidance, careful mentoring, and warm supports during the course of my Ph.D. candidature. Their broad horizon, clear thinking and enthusiasm make me learn more. Without whom I could not have fulfilled my task of research. Professor Gay's positive outlook and confidence in my research has always inspired and spurred me to explore new frontiers. I also benefited greatly from his broad contacts where exciting research ideas are exchanged with numerous elite professors from all over the globe.

I would like to sincerely thank Professor Agnes Chang Shook Cheong for her valuable advices on the pedagogical aspects of learning. Besides advising me on my research work, she also gave me valuable and constructive advices beyond research matters.

My special gratitude also goes to Dr Shen Zhiqi who provided me valuable comments on my papers. The sharing of research experience and knowledge is gratefully acknowledged.

I would like to acknowledge the research opportunity granted by Nanyang Technological University and the Information Communication Institute of Singapore (ICIS).

I am also indebted to my wife for her encouragement and support for me in my work throughout the period. Her comforting presence saw me through this challenging period. I would like to give my heartfelt thanks to my mum, my sister, my brother-in-law, my little nephew whose silent love enabled me to complete this research work.

Last but not least, I would thank all the other research students and the technicians in ICIS for their kind help and encouragement in my research work. I would also like to extend my thanks to my gymnasium mates who have individually encouraged me in one way or another throughout these years.

Summary

Economic, social and technological forces have continued to change the landscape of education and the way of learning and teaching. The critical importance of human capital, the rapid obsolescence of knowledge and training, the need for just-in-time training delivery, advancing technology, and the search for cost-effective ways to meet the diverse learning needs of learners have called for a redefinition of the traditional processes that underlie design, development and delivery of training and education [Urdan, 2000]. In addition, we can also relate the pressing need for educational reform to the widening skills gap, demographic changes [Urdan, 2000], mismatch of learning solution to learner query and demand for flexible and personalized access of life-long learning. In response to these demands, the term, e-learning, has been coined to characterize such learning innovation and transformation [Wentling, 2000].

E-learning has proven its worth by transcending both geographical and time boundaries to provide training and education to students minus the chore of providing classroom facilities or learning under time constraint. Aside from issues such as expenses and access, e-learning also offers collaborative benefits that are not easily attained in a traditional classroom setting. For example, e-learning allows students from diverse locations and cultures to exchange ideas and learn from each other. Furthermore, the interactive and multimedia nature of the learning content also attempts to provide an enhanced learning experience.

Beyond the technological benefits, the nature of the e-learning materials (i.e. size, scope and structure of the learning materials) also permits better pedagogical considerations to be implemented. For example, most e-learning designers favor designing the learning materials in smaller chunk. These smaller-grained learning materials can be pieced together or re-sequenced easily to form a new learning module to support different pedagogical approaches (i.e. support different dominant learning style). Other e-learning advantages include:

- Learning at the learner's pace
- More interactivity with the learner
- Greater relevance

While the many advantages that e-learning brings have made it fast becoming a proxy for any training associated with advancing performance, better return of investments, knowledge and skill impartation, much still need to be done at the receiving end of the e-learning initiative – **the catering to the student's needs**. Currently, even with the exploitation of technology for learning, there is a significant gap between what the learners need and what we can currently provide. Personalizing instructions to individual learner's needs is currently difficult, if not impossible, under our conventional, teacher-centric, and one-size-fits-all transmission teaching approach. While e-learning has brought new dimension to learning, its current modes of learning, unfortunately, provide little individualization.

Hence, in this thesis, we attempt to realize the learning personalization aspect of e-learning. Through an elaborate process of (1) eliciting what the learner wants to learn, (2) eliciting the learner's prior expertise in that domain of interest, (3) dynamically assembling a learning course that starts off from the learner's prior knowledge and (4) personalizing the learning instructions according to the learner's constraints and learning preferences, we propose a systematic methodology and mathematical model to provide learning personalization. Besides providing a learning personalization methodology, we also devise a pedagogical content development methodology. This pedagogical methodology, designed in collaboration with a pedagogy expert from the National Institute of Education, Singapore, lays the pedagogical foundation for structuring, designing, developing, and implementing the e-learning resources. Using our developed Goal Driven Learning Model, we show how the learner's needs can be characterized into goals. Based on the learning goals, the model determines what parts of prior knowledge are to be presented, in what sequence, and how the delivering of the desired knowledge can be supported.

A prototype system is also designed and implemented to present the research contributions and to demonstrate how the proposed theories and methodologies can be realistically implemented. Learning scenario and case studies are presented to put together all aspects of the research work and to present a learning environment that truly provides learning personalization.

This research claims the following contributions to the state of art of e-learning.

- i. A characterization of e-learning goals that is essential to personalize learning
- ii. A novel learning methodology to externalize the learner's existing learning abilities using concept mapping techniques
- iii. An extension and enhancement of the Goal-oriented modeling method to model the learning environment accurately
- iv. A novel e-learning model for realizing personalized and learner-centric learning
- v. A novel personalized algorithm and pedagogical approach for e-learning
- vi. Case studies and scenarios to explore real world e-learning domain needs. The results of the case studies aim to prove that the proposed theory and methodology are not only promising but practical.

A total of 19 papers (8 journal and transaction articles, 1 magazine article, and 10 conference articles) were published.

Content Page

1. Introduction.....	1
1.1 Introduction to Thesis.....	2
1.2 Motivation.....	3
1.3 Objectives.....	6
1.4 Approach.....	7
1.4.1 Research Problem Definition.....	7
1.4.2 Research Direction.....	7
1.5 Major Contributions of Research.....	8
1.6 Organization of Thesis.....	9
 2. Literature Review – Theoretical Background, Research Focus & Related Work.....	 10
2.1 Introduction to E-learning.....	11
2.1.1 Review of Common E-learning Definitions.....	12
2.1.2 Review of E-learning History.....	13
2.1.3 General Concepts and Terminologies Employed.....	15
2.1.3.1 Online Learning.....	15
2.1.3.2 Web-Based Learning.....	15
2.1.3.3 Distance Learning.....	15
2.1.3.4 Collaborative Learning.....	16
2.1.4 Comparison of Terminologies.....	16
2.1.4.1 E-learning Vs Online Learning.....	17
2.1.4.2 E-learning Vs Distance Learning.....	17
2.1.4.3 E-learning Vs Web-Based Learning.....	18
2.1.4.4 E-learning Vs Collaborative Learning.....	18
2.1.5 A look at E-learning Market Value.....	18
2.1.6 A look at Major E-learning Providers.....	19
2.1.7 A look at existing Research Direction.....	20
2.1.8 Nature of Learning.....	21
2.1.8.1 Behaviorism.....	22
2.1.8.2 Cognitivism.....	24
2.1.8.3 Constructivism.....	26
2.2 Research Work.....	27
2.2.1 Research Direction.....	28
2.2.1.1 Technology Concerns Precede Pedagogical Concerns.....	28
2.2.1.2 Instructor's Needs Precede Learner's Needs.....	31
2.2.1.3 Distribution Concerns Precede Learning Concerns.....	33
2.2.2 Key Research Issues.....	34
2.3 Review of Related Work.....	36
2.3.1 Importance of Learning Personalization.....	36
2.3.2 Learning Personalization Status.....	40
2.3.2.1 Current Learning Personalization Scenario.....	41
2.3.2.2 Learning Problems Identified.....	43
2.3.3 Overview of Existing Learning Personalization Approaches.....	43
2.3.3.1 Search Engines.....	43

2.3.3.2 Recommender Systems.....	45
2.3.3.3 Adaptive Hypermedia Systems (AHS).....	46
2.3.4 Review of Existing Learning Personalization Approaches.....	47
2.4 Summary.....	50
3. Learning Personalization – A pedagogical Content Development Methodology.....	52
3.1 Pedagogy Considerations for E-learning.....	53
3.1.1 Research Contribution – from Pedagogy Perspective.....	53
3.1.1.1 Gaps in E-learning Pedagogy Adoption.....	53
3.1.1.2 Contributions from Pedagogical Research.....	55
3.1.2 Learning Theories Adopted.....	55
3.1.2.1 Implications of Behaviorism on E-learning.....	56
3.1.2.1.1 Theorems to be Adopted.....	56
3.1.2.1.2 Examples in an E-learning Context.....	57
3.1.2.1.3 Pitfalls to Avoid.....	57
3.1.2.2 Implications of Cognitivism on E-learning.....	57
3.1.2.2.1 Theorems to be Adopted.....	57
3.1.2.2.2 Examples in an E-learning Context.....	58
3.1.2.2.3 Pitfalls to Avoid.....	58
3.1.2.3 Implications of Constructivism on E-learning.....	58
3.1.2.3.1 Theorems to be Adopted.....	58
3.1.2.3.2 Examples in an E-learning Context.....	59
3.1.2.3.3 Pitfalls to Avoid.....	59
3.2 Learning Resource Structure.....	59
3.3 Proposed Content Development Approach.....	61
3.3.1 Overview of Approach.....	61
3.3.1.1 Approach and Justification – ADDIE Model.....	62
3.3.1.2 Approach and Justification – Instructional Strategies (Micro-Level).....	63
3.3.1.3 Approach and Justification – Learning Events Classification (Macro-Level).....	63
3.3.1.4 Approach and Justification – Scaffolding Framework.....	64
3.3.2 Design Tasks.....	67
3.3.2.1 Instructional Strategies – 4Ts.....	67
3.3.2.1.1 4-Ts of a learning session: Target.....	67
3.3.2.1.2 4-Ts of a learning session: Training.....	69
3.3.2.1.2.1 Training Tasks.....	71
3.3.2.1.2.2 Training Materials.....	71
3.3.2.1.2.3 Training Support.....	72
3.3.2.1.3 4-Ts of a learning session: Transfer.....	72
3.3.2.1.4 4-Ts of a learning session: Transformation.....	74
3.3.2.2 Learning Events Classification – Educational and Performance.....	74
3.3.2.2.1 Educational Learning Events.....	74
3.3.2.2.1.1 Factual Learning.....	74
3.3.2.2.1.2 Conceptual Learning.....	75

3.3.2.2.1.3 Supervised Learning.....	76
3.3.2.2.2 Performance Learning Events.....	76
3.3.2.2.2.1 Guidelines.....	77
3.3.2.2.2.2 Reference.....	77
3.3.2.2.2.3 Troubleshooting.....	78
3.3.2.3 Scaffolding Framework – Design Tasks.....	78
3.3.2.3.1 Scaffolding Assessment.....	80
3.3.2.3.1.1 Justification of Scaffolding Assessment Levels.....	80
3.3.2.3.2 Scaffolding Principles.....	81
3.4 Proposed Content Development Methodology.....	82
3.4.1 Analysis Phase.....	83
3.4.2 Design Phase.....	83
3.4.3 Development Phase.....	84
3.4.4 Implementation Phase.....	85
3.4.5 Evaluation Phase.....	85
3.5 Summary.....	85
4. Learning Personalization Methodology – Theorems, Algorithms, and Mathematical Model.....	87
4.1 Learning Personalization Concepts.....	88
4.1.1 Concepts – Learning Personalization.....	88
4.1.2 Concepts – Learning Dependency.....	88
4.1.3 Concepts – Expertise Gap Identification.....	89
4.1.4 Concepts – Cognitive Map (C-Map).....	90
4.1.5 Concepts – Course Expertise Map (E-Map).....	90
4.1.6 Concepts – Personalized Map (P-Map).....	91
4.2 Learning Personalization Scenario Revisit.....	91
4.3 The Algorithms.....	92
4.3.1 C-Map.....	92
4.3.1.1 Expertise Population Rules.....	93
4.3.1.2 CMP-Algorithm 1 – CMPA1.....	94
4.3.1.3 CMP-Algorithm 2 – CMPA2.....	95
4.3.2 E-Map.....	98
4.3.2.1 Experimental Setup.....	100
4.3.2.1.1 Experimental Procedure.....	101
4.3.2.1.2 Experimental Key Processes Explained.....	102
4.3.2.1.3 Experimental Assumptions/Simplifications.....	105
4.3.2.1.4 Experimental Data.....	105
4.3.2.1.5 Experimental Sample.....	106
4.3.2.1.6 Experimental Results.....	107
4.3.2.1.7 Experimental Results Analysis.....	108
4.3.2.2 Hierarchical Weightage Considerations.....	110
4.3.2.3 Ranking Procedure.....	114
4.3.2.4 EMP-Algorithm – EMPA.....	117
4.3.3 P-Map.....	117

4.3.3.1 PMP-Algorithm – PMPA.....	118
4.3.3.2 Breadth-First Traversal Algorithm.....	119
4.4 Summary.....	120
5. Goal Driven Learning Model.....	122
5.1 Implementation Logics.....	123
5.1.1 Agent Technology.....	123
5.1.2 Agent Modeling.....	124
5.1.3 Why Goal Driven Learning.....	126
5.2 Nature of Goal Driven Learning.....	128
5.2.1 Active Learning.....	128
5.2.2 Opportunistic Learning.....	128
5.2.3 Multi-Strategy Learning.....	129
5.2.4 Realistic Learning.....	129
5.3 Goal Net Model.....	130
5.3.1 Introduction to Goal Net Methodology.....	130
5.3.2 Goal Net Model.....	130
5.3.3 Extension of Goal Net's Capabilities.....	133
5.3.3.1 Clarification between State and Goal.....	133
5.3.3.2 Clarification on State.....	134
5.3.3.3 No Mathematical Model to Model Partial Goal Achievement...	134
5.3.3.4 Extending the Graphical Notations.....	134
5.3.3.5 More information required to trigger a transition firing.....	136
5.4 Proposed Goal Driven Learning Methodology.....	136
5.4.1 Terminologies Defined.....	136
5.4.1.1 Goal.....	137
5.4.1.2 State.....	138
5.4.1.2.1 State Internal Components.....	139
5.4.1.3 Transition.....	141
5.4.1.4 Task.....	142
5.4.1.4.1 Task List.....	142
5.4.1.4.2 Check List.....	142
5.4.1.4.3 Recursive List.....	142
5.4.2 Graphical Notations Defined.....	143
5.4.2.1 State.....	143
5.4.2.2 State Granularity.....	143
5.4.2.2.1 Atomic State.....	143
5.4.2.2.2 Composite State.....	143
5.4.2.3 Types of State Status.....	144
5.4.2.3.1 Focal Status	144
5.4.2.3.2 Optional Status	144
5.4.2.3.3 Mandatory Status	145
5.4.2.3.4 Compensating Status	145
5.4.2.3.5 Choice Status	146
5.4.2.3.6 Recursive Status.....	146
5.4.3 Processing Power of Task List.....	146

5.4.3.1 Active Task List.....	146
5.4.3.2 Passive Task List.....	147
5.4.4 Types of State Transition.....	147
5.4.4.1 Direct Transition.....	148
5.4.4.2 Conditional Transition.....	148
5.4.4.3 Probabilistic Transition.....	149
5.4.4.4 Recursive Transition.....	150
5.4.4.5 Jump Transition.....	150
5.4.4.5.1 Jump-Terminate Transition.....	151
5.4.4.5.2 Jump-Continuous Transition.....	151
5.4.5 Types of Flow.....	152
5.4.5.1 Synchronous Flow.....	152
5.4.5.2 Asynchronous Flow.....	153
5.4.5.3 Choice Flow.....	153
5.5 Justifications and Applications of Modeling Approach.....	154
5.5.1 Justification of Approach.....	154
5.5.2 Common Teaching Strategies.....	155
5.5.3 Methodology Application.....	157
5.5.3.1 Methodology Step-Through.....	157
5.5.3.2 Scenario Revisit.....	158
5.5.4 Goal Net Solution.....	159
5.6 Goal Driven Learning Model Contributions.....	161
5.7 Summary.....	163
6. Framework and Prototype Design.....	164
6.1 Architectural Considerations.....	165
6.2 Framework Design.....	166
6.2.1 Functional Layers.....	166
6.2.2.1 Wire Protocol Layer.....	166
6.2.2.2 Description Layer.....	167
6.2.2.3 Discovery Layer.....	169
6.2.2 Proposed Harnessed Framework.....	170
6.2.2.1 Ontology Services.....	170
6.2.2.2 Inference Services.....	172
6.2.2.3 Agent Services.....	173
6.2.2.4 Collaborative Authoring Services.....	173
6.2.2.5 Augment Services.....	174
6.3 Prototype Design.....	174
6.3.1 E-Grid.....	175
6.3.1.1 Motivation and Proof of Concept – Harnessed Framework.....	175
6.3.1.2 E-Grid Overview.....	175
6.3.1.3 E-Grid Architecture.....	175
6.3.1.4 E-Grid Modules.....	178
6.3.1.4.1 Infrastructure Layer.....	178
6.3.1.4.2 Web Services Layer.....	178
6.3.1.4.3 Grid Middleware.....	178

6.3.1.4.4 Content Repository.....	179
6.3.1.5 Implementation Packages.....	179
6.3.1.5.1 General Retrieval Package.....	180
6.3.1.5.2 Customized Service Package.....	180
6.3.1.5.3 Basic Analysis Package.....	180
6.3.1.5.4 Advanced Analysis Package.....	180
6.3.1.5.5 Database Administration Package.....	181
6.3.2 Content Management System.....	181
6.3.2.1 Object Authoring Module (OAM).....	183
6.3.2.1.1 Learning Resource Structure.....	183
6.3.2.1.2 Conceptual Data Modeling.....	184
6.3.2.1.3 Course Description Language (CDL).....	185
6.3.2.2 Semantic Search Module.....	188
6.3.2.2.1 Basic Search.....	188
6.3.2.2.2 Advanced Search.....	191
6.3.2.3 Dynamic Concept Map Module.....	193
6.3.2.3.1 Concept Map Format (SVG).....	194
6.3.2.3.2 Concept Map Editor (.cmap).....	195
6.3.2.4 Learner Profile Management Module.....	196
6.3.2.4.1 Personal Profile Dimension.....	196
6.3.2.4.2 Cognitive Dimension.....	197
6.3.2.4.3 Learning Preferences Dimension.....	197
6.3.2.4.3.1 Preferred Learning Styles Modeling.....	197
6.3.2.4.3.2 Preferred Learning Settings.....	201
6.3.2.4.3.3 Learning Constraints.....	203
6.4 Summary.....	203
7. Typical Learning Scenario.....	204
7.1 Learning Problems Rectified.....	205
7.2 Scenario Revisit.....	207
7.3 Contributions.....	213
7.4 Summary.....	214
8. Conclusion and Future Work.....	215
8.1 Conclusion.....	216
8.2 Contributions.....	217
8.3 Future Work.....	219
8.3.1 Service-Oriented Learning Resources Structure.....	219
Publications.....	221
References.....	224
Appendix A.....	249
Appendix B.....	270
Appendix C.....	281
Appendix D.....	287

List of Figures

Figure 4.1 Expertise Gap.....	89
Figure 5.1 The basic objects of the composite state goal model.....	130
Figure 5.2 Sequence Relationship.....	132
Figure 5.3 Concurrency Relationship.....	132
Figure 5.4 Choice Relationship.....	132
Figure 5.5 Synchronization Relationship.....	132
Figure 5.6 Types of Transitions.....	132
Figure 5.7 Types of Arcs.....	132
Figure 5.8 Ambiguity in current Goal Net Model.....	135
Figure 5.9 Actual State Representation.....	143
Figure 5.10 Simplified State Representation.....	143
Figure 5.11 Atomic State.....	143
Figure 5.12 Composite State.....	143
Figure 5.13 Focal State.....	144
Figure 5.14 Optional State.....	144
Figure 5.15 Mandatory State.....	145
Figure 5.16 Compensating Mandatory State.....	145
Figure 5.17 Compensating Focal State.....	145
Figure 5.18 Choice State.....	146
Figure 5.19 Recursive-Start State.....	146
Figure 5.20 Recursive-End State.....	146
Figure 5.21 Active Task List.....	146
Figure 5.22 Passive Task List.....	147
Figure 5.23 Direct Transition.....	148
Figure 5.24 Conditional Transition.....	148
Figure 5.25 Probabilistic Transition.....	149
Figure 5.26 Recursive Transition.....	150
Figure 5.27 Jump _t Transition.....	151
Figure 5.28 Jump _c Transition.....	151
Figure 5.29 Example of Synchronous Flow.....	153
Figure 5.30 Example of Concurrency Flow.....	154
Figure 5.31 Core Teaching Strategies.....	156
Figure 5.32 Pre-Assessment Strategy.....	157
Figure 5.33 Partial Training Goal Net Solution.....	159
Figure 5.34 S ₃₁ Goal Net (level 3).....	160
Figure 6.1 Layered Architecture.....	166
Figure 6.2 Functional Layers.....	166
Figure 6.3 Proposed Harnessed Framework.....	170
Figure 6.4 ADDIE Model.....	173
Figure 6.5 Content Authoring, Assembly, Management and Delivery Process.....	174
Figure 6.6 Grid Middleware.....	176
Figure 6.7 E-Grid Infrastructure.....	176
Figure 6.8 Learning Management System Information Page.....	177
Figure 6.9 E-Grid Architecture (Implemented).....	177

Figure 6.10 Implemented Architecture Vs Proposed Architecture.....	178
Figure 6.11 Implemented Packages for E-Grid.....	180
Figure 6.12 Overview of Prototype System.....	182
Figure 6.13 Internal Modules of Prototype System.....	182
Figure 6.14 Screen Capture of Database.....	185
Figure 6.15 CDL <course> tag.....	185
Figure 6.16 Screen Capture of Create New Course Form (1).....	186
Figure 6.17 CDL <information> tag.....	186
Figure 6.18 Screen Capture of Create New Course Form (2).....	187
Figure 6.19 CDL <structure> tag.....	187
Figure 6.20 Screen Capture of Create New Course Form (3).....	188
Figure 6.21 CDL <Concepts> tag.....	188
Figure 6.22 Screen Capture of Basic Search.....	188
Figure 6.23 Screen Capture of Basic Search with error.....	189
Figure 6.24 Screen Capture of Basic Search Query.....	190
Figure 6.25 Screen Capture of Basic Search Result.....	190
Figure 6.26 Screen Capture of Advanced Search.....	191
Figure 6.27 Screen Capture of Advanced Search Result.....	192
Figure 6.28 Concept Map in SVG format.....	194
Figure 6.29 Hovering Effect.....	195
Figure 6.30 Course Concept Map Editor Environment.....	195
Figure 6.31 Editor Environment.....	196
Figure 6.32 Learner Registration Form.....	196
Figure 6.33 Screen Capture of Questionnaire.....	201
Figure 7.1 Partial Goal Net Solution.....	208
Figure 7.2 Cognitive Map (C-Map).....	209
Figure 7.3 Course Expertise Map (E-Map).....	209
Figure 7.4 Personalized Map (P-Map).....	211

List of Tables

Table 1.1 Strategic Thrusts in iN2015.....	6
Table 2.1 Gap between the Ideal E-learning Methodology and Current Practice.....	35
Table 2.2 Key Characteristics of Personalized E-learning System and their merits...	40
Table 3.1 Scaffolding Framework – Design Tasks.....	80
Table 3.2 Scaffolding Assessment Levels.....	80
Table 3.3 Analysis Phase.....	83
Table 3.4 Design Phase.....	84
Table 3.5 Development Phase.....	85
Table 3.6 Implementation Phase.....	85
Table 3.7 Evaluation Phase.....	85
Table 4.1 Section of common stopwords used in the experiment.....	103
Table 4.2 Section of domain specific stopwords used in the experiment.....	103
Table 4.3 Experimental Data.....	106
Table 4.4 Course Sample: An Introduction to Software Architecture.....	107
Table 4.5 $\text{sim}(\text{Vector}_{\text{freq}}, \text{Vector}_{\text{KeyCon}})$ Value.....	108
Table 4.6 Experimental Result Percentile.....	108
Table 5.1 Common E-learning Strategies and their Objectives.....	155
Table 6.1 Relational Database.....	184
Table 6.2 ILS Model Comparison.....	198

List of Algorithms

Algorithm 4.1 CMPA1.....	95
Algorithm 4.2 CMPA2.....	98
Algorithm 4.3 EMPA.....	117
Algorithm 4.4 PMPA.....	119
Algorithm 6.1 ILS ACT/REF Counting Algorithm.....	199
Algorithm 6.2 ILS Scoring Algorithm.....	200
Algorithm 6.3 ILS Rating Algorithm.....	201

List of Flowcharts

Flowchart 4.1 Basic Search Retrieval Procedure.....	116
Flowchart 4.2 Advanced Search Retrieval Procedure.....	117

List of Listings

List 4.1 Expertise Population Rules.....	93
List 4.2 Procedure for Basic Search Retrieval.....	115
List 4.3 Procedure for Advanced Search Retrieval.....	116
List 4.4 Procedure for Breadth-First Traversal.....	120

CHAPTER 1

INTRODUCTION

This chapter introduces the field of research and presents the problem definition on which the research is based. This chapter consists of six sections. It starts off by introducing e-learning in section 1.1. Next, section 1.2 presents our research motivation. Besides basing our research motivation on our literature findings, we also base it on facts presented in the recently launched Singapore Masterplan iN2015 (Singapore Intelligent Nation 2015) and the Singapore National Library Board's (NLB) Library 2010 (L2010). The main goal of the research presented in this dissertation is highlighted in Section 1.3. Section 1.4 presents the research approach. Specifically, the problem statement is defined. The problem statement states the gap between the desired situation and the existing e-learning situation from the learning point of view. This statement limits and explains the research focus and direction. The research contribution is presented in Section 1.5. This chapter is concluded in section 1.6 with an overview of the thesis organization.

1.1 Introduction to Thesis

The importance of education is clear. Education is about preparing people for the future. It is about the acquiring of knowledge to put one's potentials to maximum use so that one is able to receive, comprehend and react to information from the external world. Education has always been a key priority in Singapore's nation building as reflected in its investment in education. For the year 2004/2005, the public expenditure on education was 3.5 per cent of Gross Domestic Product (GDP), or about S\$6 billion. This compares with the average expenditure of 5.4 per cent of GDP spent in developing countries in the Organization for Economic Co-Operation and Development (OECD) (Source: Education At A Glance Report, OECD indicators, 2005). Over the years, Singapore has also invested well and has built a strong education system, which is a source of national competitiveness [iN2015, 2006].

With the rapid advancement of the technology and the information highway, it is important to note that the face and nature of education is changing. While the fundamentals of learning and teaching remain relatively unchanged, the learners of today have changed radically. Today's students are no longer the people our traditional educational system was designed to teach. The arrival and rapid dissemination of digital technology in the last decades of the 20th century have not only permeated the everyday life of today's learners but also provided the impetus for the exploitation of Information Technology (IT) for education.

As information technology becomes more robust, easier to use and affordable, it increasingly permeates learning activities. New and exciting learning opportunities are inspiring teachers and motivating the learners. Beyond reading from printed learning materials, learners can also learn through interactive games, virtual scenarios and simulations. Beyond face-to-face consultation with teachers, learners can also work collaboratively with peers or be supported by online scaffolding aids. Beyond the constraints of time and space boundaries, learners can transcend traditional limitations to communicate, learn and interact with peers across the continents. Beyond the confines of classroom and curriculum schedules, learners can learn where their interest takes them. Beyond one instructional teaching and learning strategy for all, teachers can customize

learning contents according to the learning needs, preferences and abilities of students. Beyond learning in a confined space such as school, the learners can form virtual online communities and share with like-minded individuals.

The exploitation of Information Technology (IT) for education, commonly defined as e-learning, is fast shaping the way we learn as well as our needs for learning. However, e-learning does not concern only the academic sector. E-learning is also fast gaining importance in the corporate sector. Many in the corporate sector are expecting e-learning to become a business critical issue for companies and see e-learning as an important instrument for human resource development [De Vries, 2005]. As the speed of change in technology and the general business environment continues to accelerate, corporate companies are also taking advantage of the new Net-based infrastructure for learning to maximize and gain competitive success.

1.2 Motivation

One of the most intriguing challenges in education – be it for the academic educators, students, or for corporate employers, employees and training providers – is how to impart knowledge effectively. From the students' perspective, their main concern is whether they are learning anything that adds to their repertoire of knowledge, while for employees, the main concern is whether the training helps in their work. Looking from the employers' perspective, if they are funding the training, it is only fair if a certain amount of 'tangible' return of investment is evident. As for the educators and training providers, they are concerned about how much knowledge their students are acquiring in respect to the long hours that they have invested on the training materials. Hence, learning/teaching effectively is essentially the concern for all.

Acknowledging the importance of imparting knowledge effectively, the next question to ask is: How can we impart knowledge effectively? Before we can answer that, we need to take a step back and think: How can we know what to teach? With the proliferation of information, access to information is never a problem. Literally, the Internet has expanded access to information exponentially. As Chee [Chee, 2004] puts it, *"if access is all that matters, education would be reduced to a trivial matter of simply providing an*

excellent library then all the learners will get an educated citizenry”. Hence, this research advocates that the crucial part in our attempt to impart knowledge effectively lies not in the provision of information access but on how to provide personalized learning materials. Specifically, four crucial tasks are essential. They are the elicitation of (1) what the learner wants to learn, (2) what is the learner’s prior expertise in that domain of interest, (3) how to dynamically assemble the learning course that starts off from one’s prior knowledge and (4) how to personalize the learning instructions according to the learner’s constraints and learning preferences.

However, it is difficult, if not impossible, to effectively provide personalized learning materials under current conventional, teacher-centric, and one-size-fits-all transmission teaching approach. Currently, our educational system is too rigid. We often fail to understand (or rather, we choose to ignore) that students are DIFFERENT, with different backgrounds, knowledge, interests and learning styles. Each student should be treated individually. But our current modes of learning provide little individualization. Every student tends to be provided with the same learning experience. While this approach to learning works for some, many do not learn, or learn only partially. The major learning modes in schools are still lectures and textbooks which provide little individualization.

While the employment of technology to aid education has made a huge impact, it is still very much of a promise unfulfilled. Technology has without a doubt brought new dimensions to learning. However, most online information systems are still directly translating existing curriculum materials into online information (excepting inclusion of sophisticated multimedia design which is prohibitively expensive). Hence, it is not too surprising to find that online materials are also following the curriculum approach of treating all students equally.

As this research stresses, before we can truly impart knowledge effectively, we first need to have an educational system that is able to individualize learning materials to the needs of each student. The need to provide learning personalization is also being urgently promoted in Singapore. This is reflected by the recently launched Masterplan, Singapore Intelligent Nation 2015 (iN2015) and Singapore National Library Board’s (NLB) Library

2010 (L2010), which aims to utilize infocomm to deliver personalized learning and services.

Singapore NLB Library 2010 (L2010) was unveiled on 21st July 2005. L2010 is NLB's second Masterplan to put forwards NLB's strategic direction for the next five years. It aims to make the public libraries in Singapore, pioneers in the use of infocomm to deliver personalized library services. The L2010 plan is NLB's response to the changing needs of Singapore's society and economy and envisions bringing the world's knowledge to Singapore to create positive social and economic impact. Specifically, in order to deliver personalized library services, the L2010 plan will focus on these five key areas:

- Building a network of knowledge assets and making them accessible;
- Leveraging on technology especially in support of collaboration;
- Organizing customer communities to serve them better;
- Expanding its professional competencies to deliver on its objectives; and
- Measuring its impact to ensure continuing value and relevance.

The Singapore Intelligent Nation 2015 (iN2015) Masterplan was launched in 20th June 2006. iN2015 is a blueprint to navigate Singapore's exhilarating transition into a global city that is universally recognized as an enviable synthesis of technology, infrastructure, enterprise and manpower. iN2015 plans to provide seamless access to intelligent technology to connect, innovate, personalize and create. Led by the Infocomm Development Authority of Singapore (IDA), iN2015 is Singapore's 10-year Masterplan to realize the potential of infocomm over the next decade. It is a multi-agency effort that is the result of private, public and people sector co-creation.

iN2015's main goal is to foster an engaging learning experience to meet the diverse needs of learners in Singapore, through the innovative use of infocomm. In order to realize this goal, the education and learning community identifies learning personalization as one of the three key strategy thrusts that will make the largest impact.

Strategic Thrust 1	Strategic Thrust 2	Strategic Thrust 3
Creating an enriching and personalized learner-centric	Building a nation-wide Education and Learning infrastructure	Positioning Singapore as a centre for innovation in the use of

environment in our educational institutions		infocomm technologies for the Education and Learning Sector
<ul style="list-style-type: none"> • Use infocomm to support changes in pedagogies in our educational institutions • Develop new learning resources and new infocomm-enabled assessment modes • Build capabilities of teachers, school leaders and curriculum planners • Develop incubator educational institutions that will generate innovation in the use of infocomm to support engaged learning 	<ul style="list-style-type: none"> • Make broadband infrastructure affordable and accessible to educational institutions • Build a network of knowledge assets for lifelong learning 	<ul style="list-style-type: none"> • Forge strategic partnerships with key companies and research institutions in this field, and locate test-bedding, prototyping and R&D centers in Singapore • Develop a R&D agenda on new technologies and models for harnessing infocomm in E&L • Develop capability in industry to harness infocomm for E&L

Table 1.1 Strategic Thrusts in iN2015 (Source: iN2015 Education and Learning Report)

Our research vision of learning personalization is in line with the iN2015 vision that states that:

“Learners in 2015 will access the latest learning resources using personalized learning devices.... Learners can choose to learn anytime and at any place Educators guide learners, by customizing learning plans and resources, and using new assessment tools to monitor their progress....”

iN2015 Education and Learning Report [iN2015, 2006]

However, while learning personalization is crucial, little such material exists. But there is enough evidence to show that we can prepare such material. The advancement of the information and communication technologies coupled with the interactive capabilities of modern computers has laid down a great foundation for research to probe into how we can investigate what the learner needs and also to provide suitable individualized help.

1.3 Objectives

- i. To provide an e-learning model that truly provides (1) personalized and (2) learner-centric learning
- ii. To explore and extend a Goal oriented modeling method for modeling and governing e-learning goals

- iii. To include sound learning pedagogy (scaffolding learning activities and mediation of learning experiences) as a process model to govern the learning and teaching process
- iv. To demonstrate and put into practice the proposed methodologies through a series of test cases set in a real world e-learning domain
- v. To make recommendations on further research studies that are needed to enhance the proposed methodologies

1.4 Approach

1.4.1 Research Problem Definition

This research identifies the lack of **learning personalization support** as the main research problem. We focus on the act of learning from the learner's point of view and investigate how e-learning can provide timely, on-demand access to relevant, personalized knowledge that caters specifically to individual learning preferences, styles and prior knowledge.

[Definition 1.1] Learning personalization is defined as the process of developing individualized learning programs for each learner with the intent of engaging a learning process that is characterized by prior knowledge, current needs and learning preferences.

1.4.2 Research Direction

We address the research problem from three aspects and hope to answer these questions.

1. Pedagogy Aspects

What work does the concept of e-learning do in the traditionally proven theories of learning and instructions? How can e-learning theories, if any, stand in relation to the theory of learning? Is the proposed personalized learning methodology justifiable by existing learning theories?

2. Personalization Aspects

How does the employment of technology for e-learning, however articulated and sophisticated in multimedia design, stand in relation to the learner's specific goals in everyday cognitive, work practices and in formal and informal education settings? How can the learning content be personalized to each individual's learning abilities, preferences and constraints?

3. Implementation Aspects

When kinds of Internet-facilitated learning experience should be supported by e-learning? What are the best practices for e-learning that should be incorporated to enhance the learning experience? How can the proposed pedagogical and personalization methodologies be realized? How can teacher/trainer's goals be realistically modeled and translated into learning goals for the learners?

Our research approach in learning personalization is to provide an elaborate process of identifying learner's learning preferences, current knowledge level and learning needs at real-time before delivering a personalized learning experience that synthesizes new knowledge directly into the learner's existing 'cognitive structure'.

Specifically, the following tasks are investigated.

- Elicitation of learner's learning goal;
- Elicitation of learner's existing learning abilities (prior knowledge elicitation);
- Provision of a dynamic, on-demand learning course that starts off from the learner's prior knowledge;
- Personalizing the learning instructions according to the learner's constraints and learning preferences.

1.5 Major Contributions of Research

This research claims the following contributions to the state of art.

- i. A characterization of e-learning goals that is essential to personalize learning
- ii. A novel learning methodology to externalize the learner's existing learning abilities using concept mapping techniques
- iii. An extension and enhancement of the Goal-oriented modeling method to model the learning environment accurately
- iv. A novel e-learning model for realizing personalized and learner-centric learning
- v. A novel personalized algorithm and pedagogical approach for e-learning
- vi. Case studies and scenarios to explore real world e-learning domain needs. The results of the case studies aim to prove that the proposed theory and methodology are not only promising but practical.

1.6 Organization of Thesis

- Chapter 1 provides the reader with a background and introduction to the research topic of the thesis. It introduces the field of research and presents the problem definition on which the research is based.
- Chapter 2 introduces the literature of e-learning and learning theories to prepare the grounds for discussion of the research work. Work pertaining to the research domain, namely, learning personalization status and existing effort, is reviewed.
- Chapter 3 describes our learning personalization work from a pedagogist perspective. It introduces our proposed pedagogical content development methodology, e-learning pedagogy model and scaffolding framework. This chapter contributes to the e-learning pedagogy field and is done in collaboration with a pedagogy expert from the National Institute of Education, Singapore. Appendices A and B supplement this chapter.
- Chapter 4 provides the theoretical background for the proposed learning personalization methodology. Building upon our pedagogical framework, the formal definitions and concepts underlying the methodology are discussed. An experimental setup, procedure, results, observations and analysis are documented in this chapter to justify some of the proposed work. Appendices C and D supplement this chapter.
- Chapter 5 describes our Goal Driven Learning Model. This model allows the characterization of learner's needs into goals. Based on the learning goals, the model determines what parts of prior knowledge are to be presented, in what sequence, and how the delivering of the desired knowledge can be supported.
- Chapter 6 presents the documentation of the entire design process of the learning environment. The architectural considerations and framework is presented.
- Chapter 7 discusses the prototype system that is designed to present the research contributions and to demonstrate how the proposed theories and methodologies can be realistically implemented. A learning scenario is also presented to put together all aspects of the research work and present a learning environment that truly provides learning personalization.
- Chapter 8 concludes this thesis. The main contributions and proposed future work are presented.

CHAPTER 2

LITERATURE REVIEW – THEORETICAL BACKGROUND, RESEARCH FOCUS AND RELATED WORK

This chapter will describe the field of research and present the concerns on which the research is based. It consists of four sections. The chapter starts off by introducing the literature of e-learning to prepare the grounds for the discussion of our research work. Next, section 2.2 presents our research work. Specifically, the research direction, concerns and key research issues are defined. The gap between the desired situation and the existing e-learning situation from the learning point of view is identified. This explains the scope of the research focus and contributions. Section 2.3 reviews related work to our research domain – learning personalization status and existing learning personalization efforts. As an illustration of the current limitations in learning personalization efforts, a typical learning scenario, based on the Singapore Educational System and follows the Ministry of Education learning guidelines (see <http://www.moe.gov.sg/>) is presented. This scenario is employed to motivate the thesis with a concrete application. Subsequently, this scenario will be revisited throughout this thesis to illustrate the theories, algorithms and methodologies that are presented in the later chapters. This serves as concrete examples that come directly from the application of our proposed solutions. This chapter is summarized in section 2.4 with our vision for future e-learning research and development.

2.1 Introduction to E-learning

“The biggest growth in the Internet, and the area that will prove to be one of the biggest agents of change, will be e-Learning. ...

Education of the Internet is going to make e-mail usage like a rounding error in terms of the Internet capacity it will consume”

John Chambers, CEO, Cisco Systems [Chambers 1999]

E-learning, by virtue of its uniquely distributed, asynchronous nature, shows promise in fostering significant learning improvements especially in accessibility and opportunity. It couples the advancement of technology and the advent of the information highway to eliminate barriers of time, distance and socioeconomic status, thereby creating a whole new dimension to learning. With its capability of transmitting knowledge faster and more effectively, many people are accepting e-learning as a means of upgrading themselves and keeping up with the rapid changes that define the Internet world.

The e-learning paradigm shift capitalizes on two main aspects: the elimination of the barriers of time, space and distance, and the personalization of learner experience [Sampson, 2002a]. Typical demands include [Rosenberg, 2001; Sampson, 2002b]:

- **Personalization**, where instructions are customized to an individual learner based on the analysis of his needs, objectives, current status of knowledge and learning style preferences. This includes constant monitoring of progress and feedback;
- **Interactivity**, where the learners can experience and immerse themselves in a learning environment where they can control, combine, manipulate different types of media, such as text, sound, video, computer graphics, and animation;
- **Media-rich context**, where the educational materials can exist in more than one form of presentation styles;
- **Just-in time delivery**, where on-demand learning materials, support and real-time assistance can be provided at the right time to the right learner;
- **User-centric**, where the learners take control over their own learning experience.

The unique characteristics of e-learning technology (i.e. the distributed nature of the distance learning modality, the physical separation of learners from instructors, the

asynchronous communication paradigm, etc.) however, require adjustments in the nature of instruction specifically designed for that modality. The development of instructional tools and methodologies to assist in this process, although not a new concept, remains of continuous interest to much of the educational technology community. It is in this vein that this research sets out to investigate novel research methodologies to realize e-learning's potential as an imperative supplement and enhancement of traditional methods.

Following this introduction, a review of common e-learning definitions, its history, general concepts, terminologies and their distinction lays the theoretical foundation for the research. The discussion on major e-learning providers, existing research direction and its estimated market value is used to justify the research value of e-learning and our research work in particular. Lastly, the discussion on the nature of learning completes our review of e-learning.

2.1.1 Review of Common E-learning Definitions

To date, e-learning has a plethora of definitions. However, none of these definitions have won overwhelming acceptance. The lack of an agreed upon definition is also due to the fact that e-learning has evolved from the earlier concepts such as TBT (technology-based training) and CBT (computer-based training). As these concepts have no agreed upon definition, the chance that e-learning would or will is pretty remote. Moreover, as e-learning is widely adopted [Manjunath, 2006] and spans different fields of application (academic, corporate, research, etc.), each different group of stakeholders within their own institutional context, are describing e-learning very differently. The overlays of technology add on another challenge to reach a common understanding because the technicalities are often unfamiliar to many of the stakeholders whose fields of expertise are usually not technological in nature [OCLC, 2003].

However, there are some notable definitions and they are listed below:

1. E-learning is simply '*using the Internet as a communications medium where the instructor and students are separated by physical distance*' [Cooper, 1999]
2. '*E-learning is just-in-time education integrated with high velocity value chains. It is the delivery of individualized, comprehensive, dynamic learning content in real*

- time, aiding the development of communities of knowledge, linking learners and practitioners with experts*’ [Drucker, 2000]
3. E-learning is described as ‘*supporting a learning experience by either developing or applying Information & Communication Technology (ICT)*’ [Wade, 2002]
 4. E-learning is defined as “*the delivery of education (all activities relevant to instructing, teaching, and learning) through various electronic media. The electronic medium could be the Internet, intranets, extranets, satellite TV, video/audio tape, and/or CD ROM.*” [Koochang, 2005, pp. 77]
 5. E-learning is defined as “*the use of the Internet to access learning materials; to interact with the content, instructor, and other learners; and to obtain support during the learning process, in order to acquire knowledge, to construct personal meaning, and to grow from the learning experience*” [Sharifabadi, 2006]

Other definitions of e-learning have different words that are basically similar to one or more of the aforementioned definitions. Many include the Web, the Internet and/or ICT, whether explicitly or implicitly, in their definition. In fact, the ‘e’ before the learning may refer just to that when the term e-learning was coined. Taking into account these definitions, it can be surmised that all of these definitions comprise the combination and interaction of the following components: learning, communication and teaching via different electronic media.

However, as with many learning concepts that are felt to have a significant impact in theoretical and practical schemes, we believe it is worthwhile to do some historical excavation, identify and circumscribe e-learning’s early uses and uncover the roots of the concept, before analyzing its worth, future uses and direction.

2.1.2 Review of E-learning History

While e-learning is widely adopted and anchored strongly in many training strategies, the roots of e-learning are surprisingly unheard of. This research believes that the roots of e-learning stem from the earlier version of the CBT whose learning materials were the first to be stored and distributed on CD-ROMs. CBT is usually multimedia-based training [Zahm, 2000] and offers the convenience of self-paced learning [Karon, 2000].

Subsequently, the latest ICT developments present opportunities for both technological breakthroughs and theoretical advancements in the educational domain. This coupled with the multitude of resources available in the World Wide Web underscores the importance of using ICT effectively in education and catalyzes research work on improving traditional learning. Research efforts are being carried out on two fronts: Technically, the Internet together with the fast growth of computer and communication technology (namely interactive multimedia) is increasingly being used to support education and learning. Theoretically, strategies on the effective implementation and application of the many pedagogical theories from past literature are being revisited for repurposement to suit today's learners.

Along with fast advancing technologies and research efforts came digitalized learning contents and the exploration and practice of new educational paradigms. Subsequently, key transformations in traditional learning are under way: from classroom-based learning to web-based learning, from printed materials to online materials, from face-to-face interaction to virtual interaction, from fixed semester-based learning to learning-on-demand, from traditional learning to lifelong learning and from the use of physical facilities to the use of networked facilities.

As a result of these educational shifts, several new terms for educational paradigms were coined: Online learning, Web-based Learning, Distance Learning, Collaborative Learning and E-Learning. These terms are often used in a loosely interchangeable fashion, with e-learning being used as the current buzzword for modern education. The proliferation of terms has however, caused confusion even within the education community as different groups attach different meanings to the same term and concept [Morris, 2002]. Some people have argued that Web-based training is a subset of e-learning while others contend that CBT is a broad term that may include both CD-ROMS and the Internet. Occasionally, even technical literature diverges from the common usage of these terms, either over-generalizing or restricting the meaning of the terms [Tsai, 2002]. However, these terms represent concepts with subtle, yet consequential differences. Therefore, a thorough familiarity with each concept and its distinctive characteristics is a crucial factor in establishing adequate specification, evaluating its value and promise and in enabling and

promoting effective learning practices. Hence, in this research, we discuss the differences, review the governing concepts underlying the terms, compare and suggest relationships, if any, between the terms.

2.1.3 General Concepts and Terminologies Employed

2.1.3.1 Online Learning

Online learning is defined by Schreiber & Berge [Schreiber, 1998] as any technology-based learning, that is, information currently available for direct access. They further added the assumption of a linkage to a computer. Gotschall [Gotschall, 2000] agrees with this definition and classifies online learning as an all-encompassing term that refers to training that is done with a computer over a network. Although the concept of online learning predates the appearance of the Web, most publications view online learning as referring to materials delivered over the Internet or intranets [e.g. Malopinsky, 2000; Schank, 2001; PBS, 2001]. The levels of sophistication of online learning can vary. It can extend from simple text-based programs to those which include sophisticated animation, simulations, online mentoring and discussion functionalities [Urdan, 2000].

2.1.3.2 Web-Based Learning

Web-based learning is defined as a generic term for training and/or instruction delivered over the Internet or company intranet using a Web browser [Hall, 1997]. Accessibility of this training, related by Hall [Hall, 1997], is through the use of browsers such as Internet Explorer or Netscape Navigator. Web-based training typically includes static methods – such as streaming audio and video, hyperlinked Web pages, live Web broadcasts, and portals of information – and/or interactive methods – such as bulletin boards, chat rooms, instant messaging, videoconferencing and discussion threads. The instructions can be facilitated and paced by the trainer or self-directed and paced by the learners.

2.1.3.3 Distance Learning

Distance learning is a type of learning in which learners take academic courses by accessing information and communicating with the instructor asynchronously over a computer network. It is defined as a learning process that must fulfill three criteria: a geographical distance and time separates communication between the trainer and the

learner [Perraton, 1988], the volitional control of learning by the student rather than the distant instructor [Jonassen, 1992], and noncontiguous communication between student and teacher, mediated by print or some form of technology [Keegan, 1986; Garrison, 1987]. Learners can work on their own home or at the office and communicate with the faculty and other learners via e-mail, electronic forums, videoconferencing, chat rooms, bulletin boards, instant messaging or other forms of computer-based communication. Most distance learning programs often include a CBT system and communications tools to produce a virtual classroom. It is important to note that simply posting or broadcasting learning materials to (distanced) learners is not distance learning. Instructors must be involved in the learning process (i.e. receive and respond to feedbacks) [Tsai, 2002].

2.1.3.4 Collaborative Learning

Collaborative learning refers to an instructional method in which learners at various performance levels work together in small groups toward a common goal. The learners are responsible for one another's learning as well as their own. Thus, the success of one learner helps other learners to be successful. Proponents of collaborative learning claim that the active exchange of ideas within small groups not only increases interest among the participants but also promotes critical thinking. According to Johnson et al. [Johnson, 1986], there is persuasive evidence that cooperative teams achieve at higher levels of thought and retain information longer than learners who work quietly as individuals. The shared learning gives learners an opportunity to engage in discussion, take responsibility for their own learning, and thus become critical thinkers [Totten, 1991].

2.1.4 Comparison of Terminologies

While e-learning is similar in certain aspects to the other terminologies of education, e-learning has its apparent differences and should not be used interchangeably. For each of these concepts, there exist discriminating features that serve as the primary characteristics of the learning activity. Intensive use of those features is required. Incidental or occasional usage of the characteristic features is not sufficient to qualify for a certain type of learning. We believe that failing to recognize the differences between these concepts, however fine, precludes precise communication and limits the pace of research development. In fact, we feel that the failure to adequately define the terminology

employed is a reflection of a poor understanding of the available technologies. The ability to refer to each concept appropriately not only conveys precise and accurate messages, but also entails correct actions and provides a clear view of challenges, potentials, and trade-offs. In the end, recognizing such subtle differences enables faster, more accurate and discerning research.

2.1.4.1 E-learning Vs Online Learning

This research contends that e-learning represents the whole category of technology-based learning while online learning constitutes only one part of such learning that requires learning to be delivered via Internet, Intranet or Extranet. Hence online learning can be seen as a sub-set of e-learning and is synonymous with Web-based learning.

2.1.4.2 E-learning Vs Distance Learning

The general acceptance that the ‘e’ part of e-learning refers to the electronic support to the learning process has mistakenly led some people to see e-learning as a synonym to distance learning [Hentea, 2003] while others see e-learning as a subset of distance learning [Urduan, 2000]. However, we agree with Schank [Schank, 2001] that e-learning is not merely distance learning. This research advocates that distance learning is, at best, one form of e-learning. In fact, distance learning should fall under the umbrella of e-learning. Although distance learning provides the base for e-learning’s development by separating the teacher in time and place from the learner and by demonstrating that effective learning outcome can be achieved outside the traditional face-to-face educational framework of the classroom, e-learning embraces more than distance education. The concept of distance learning involves considerable amount of off-campus study, supported by either online or printed training materials sent by the distance education institution. In distance learning, the learner audience is never (or very rarely) in physical proximity to the instructor and the instructor-led traditional classroom sessions are either eliminated, adjusted for some different form of non-real time interaction, or replaced with real-time virtual classroom. While e-learning supports distance learning, its main emphasis is not on the distance aspect. Instead, e-learning focuses on a broad view of learning that goes beyond the traditional paradigms of education. While e-learning supports the use of distance learning, it can also be used as a tool to aid and supplement

face-to-face learning. Hence, it is more appropriate to view distance learning as one type of e-learning solution rather than as a different form of learning.

2.1.4.3 E-learning Vs Web-Based Learning

Much literature associates e-learning with Web-based learning over the Internet [e.g. Rosenberg, 2000; Horton, 2000; Driscoll, 2002] and used the term e-learning interchangeable with Web-based learning. However, this research argues that such conception is too narrow as there are many distributed e-learning systems that utilize the Internet but does not reside within a Web browser.

2.1.4.4 E-learning Vs Collaborative Learning

Collaborative learning perhaps is the most distinctive among the terms. A typical learning process in a collaborative learning environment requires the learners to efficiently share, stock and reuse the various learning resources and exchange heuristic knowledge generated in the collaborative learning environment [Okamoto, 2005]. The concept of collaborative learning focuses not on the information-oriented applications of the Internet technologies but on the exciting potential of a new learning medium about community and collaboration [Bruckman, 1999]. It emphasizes the creation of groups of online learners that can motivate and support one another's learning experience. While e-learning supports collaborative learning through the use of interactive forums and knowledge exchanges, it represents only one medium where learning can take place.

2.1.5 A look at E-learning Market Value

There has been much written about e-learning, definitions and opinions abound. While critics against e-learning are plentiful [Lytras, 2001], it is worthwhile to note that none has called for its demise. On the contrary, the e-learning market is expanding at an alarming rate. The increasing emergence of countless virtual universities, e-learning solution providers, ballooning value of the e-learning market (just the U.S. corporate market alone is estimated at 5 billion dollars in 2004, and expected to grow to 13.48 billion dollars by 2008 [Bounds, n.d.]) and changing knowledge landscape (i.e. new economy evolutionary processes, fast changing pace of technology, shortening product development cycles, lack of skilled personnel, competitive global economy, etc.) is

fueling its strategic importance and realization [Tim, 2000]. Also, a Google search on the terms of ‘future’ and ‘e-learning’ returned 912, 000 results in November 2003 but returns a dizzying 13, 300, 000 results in July 2006. Academically, for the last two years, Google scholar records an astounding 26 times increase in its scholarly publications (publications currently stand at 19, 900).

2.1.6 A look at Major E-learning Providers

Today, the e-learning community is a rapidly growing one that includes educational institutions, research centers, libraries, government agencies, commercial enterprises, software and hardware companies, and advocacy groups [Tynjälä, 2005]. Various new technologies and terminologies – such as learning management systems (LMS), learning content management system (LCMS), authoring and collaboration tools, synchronous and asynchronous learning platforms that connect disparate learners in virtual classrooms and streaming multimedia technologies – have been developed around the e-learning concept. A cottage industry of mostly small technology start-ups has evolved into sizable cadres of e-learning providers that include larger technology players and professional-service firms. The principal players active in e-learning can be grouped roughly into three main categories. First, providers of content – often corporate and IT training – act as the main supply of learning content. The content providers either develop the content themselves, aggregate the content developed by others or custom design content that cater to the specific needs of an organization. Some vendors also license the content from academic institutions or professional associations. Examples include SmartForce, Quisic, Litespeed Education Pte Ltd (LES), and ACE-Learning Pte Ltd. Second, providers of learning platforms provide a range of software programs that facilitate the development and delivery of e-learning courses. Such vendors claim to provide a more comprehensive learning package that ranges from pure content providers to learners’ progress monitoring and course record-keeping. Examples include Blackboard, WEBCT, WBT Systems, Lotus Development Corporation and knowledge platform. Third, providers of services offer a variety of learning-related services – learning hubs or portal companies - offer learners or organizations consolidated access to learning and training resources from multiple sources. Examples include Learn.com, Smartplanet, VCampus and EMind.

However, the lack of standardized e-learning standards and the need to gain competitive advantages in the fast-saturating market has led many e-learning players to extend their services and proclaim having a complete e-learning solution. For example, SmartForce which starts off as a content provider has subsequently merged with SkillSoft Corporation to provide a complete enterprise solution while WEBCT has also established its own e-learning hub [WebCT, n. d.].

2.1.7 A look at existing Research Direction

Research for an effective development and utilization of e-learning is likely to be a priority for many organizations in the future [Saber, 2005]. E-learning is still very much an evolving discipline and not a finished product [Angehrn, 2001]. As a matter of fact, we believe that e-learning is still far from being a finished product. In most of the cases, e-learning initiatives have not been able to keep their promise [Lytras, 2001a; Ehlers, 2004; Guri-Rosenblit, 2005].

Research efforts on e-learning have been relatively heterogeneous as e-learning is inherently multidisciplinary. Research into e-learning requires partnerships between those who develop technology and a broad range of social science researchers who seek to understand the nature of learning and the interaction and organizational effects of technology. This combines perspectives, methods and theories from the technical domains (e.g. Computer Science, Technology, Artificial Intelligence), design disciplines (e.g. Design, Human-Computer Interaction (HCI), the learning sciences (e.g. Educational Technology, Psychology, Education) and the disciplines studying communication, communities and discourse (i.e. Social Sciences, Linguistics) [Taylor, 2005]. As such, research topics for e-learning covers a wide range and have ranged from technical research issues to standardization issues to learning pedagogical issues.

Technical research aims to exploit the potential that emerges from the new technological advances to support e-learning. Some examples of such research covers design and implementation concerns that include technology acceptance modeling [Davis, 1989; Venkatesh, 2003], interactive media and presentation [Zang, 2003; Baecker, 2004], network and bandwidth delivery technologies [Greenagel, 2002; Zang, 2003] and e-

learning infrastructure [Collier, 2002]. Efforts have also been extended to mobile and ubiquitous computing [Trifonova, 2003; Trifonova, 2004; Millard, 2005a; Laroussi, n. d.], the Semantic Web [Stojanovic, 2001; Koper, 2004a], the incorporation of Grid infrastructure [Millard, 2005a], ontology [Gašević, 2005; Verbert, 2006] and concept mapping [Millard, 2005b; Tan, 2006].

Research to standardize e-learning methods, models, policies and learning content is also being carried out to encourage the uptake of e-learning facilities. More importantly, such research aims to fulfill the e-learning's vision of interoperability and reusability whereby learning contents can be truly assembled and deployed on the fly. Examples of such research cover standards on reuse [Conlan, 2002a; Malcolm, 2005], learning objects [Conlan, 2002a; Hodgins, 2002; Malcolm, 2005] and metadata and ontologies [Brase, 2003; Brasher, 2003; Currier, 2004, IEEE, n. d.].

Pedagogically, research is carried out to investigate and justify e-learning employment in terms of proven learning theories. It aims to increase the scope of e-learning to promote lifelong learning and use it to support learning outside formal educational settings. Some examples of pedagogical issues include knowledge management [Maurer, 2001; Wild, 2002] and student and faculty satisfaction and performance issues [Webster, 1997].

A large budget has been invested purely on research. For example, in the United States, the enterprises invest 20% of their budget (13 billion US-dollars) [Corporate University Exchange, 2000] while the European market of corporate e-learning has estimated its investment to grow from a capacity of 829 million dollars in 2001 to 7.4 billion dollars in 2004 [Scienter, 2001]. Unfortunately, these impressive figures have not translated into e-learning growth. It is reported that the impact of research remains limited so far, indicating that too few research results are picked up in actual practice [Wolpers, 2004].

2.1.8 Nature of Learning

The nature of learning needs to be mentioned as the way we define learning and what we believe to be the underlying learning philosophy will have important implications for e-learning. Learning theories provide e-learning researchers with verified instructional

strategies and techniques for facilitating learning as well as for setting the foundations for future research. It is the most important information and should be the governing factor that precedes all other e-learning considerations. However, many e-learning developers often operate under the constraints of a limited theoretical background or rely on the general theories of learning.

While the e-learning arena has also seen various researchers trying to create new learning theories, none has emerged as key theories in the area. Hence, the three broad learning theories (of the past), namely, behaviorism, cognitivism and constructivism, continue to remain as the most often utilized theories in the creation of instructional environments. However, it is important to note that these theories were developed at a time when learning had not yet been impacted by technology. Therefore, special considerations must be taken into account to address the limitations that were traditionally deferred due to technological constraints. Nonetheless, these learning theories provide strong structured foundations for planning and conducting instructional design activities that remain fundamental to the art of learning.

Each learning theory is discussed in terms of its specific interpretation of the learning process and the resulting implications it has for educational practitioners in general. The information presented in this section serves as an introduction to the three different viewpoints and sets the theoretical foundation for these viewpoints to be translated to our proposed learning personalization methodology (discussed in chapter four).

2.1.8.1 Behaviorism

Behaviorism, as a learning theory, is manipulative and seeks not merely to understand human behavior, but to predict and control it. It focuses on objectively observable behaviors and discounts mental activities. Behaviorism equates learning with changes in either the form or frequency of observable performance. Hence, learning, as defined by behavior theorists, is nothing more than the acquisition of new behavior and learning is accomplished when a proper response is demonstrated following the presentation of a specific environmental stimulus.

Behavior theorists view learners as being controlled by external stimuli, reinforcers, which mold behaviors by rewarding the ‘good’ behavior and punishing the ‘bad’ behavior. No attempt is made to determine the structure of a student's knowledge nor to assess which mental processes are necessary for them to use [Winn, 1990]. The learner is characterized as being reactive to conditions in the environment as opposed to taking an active role in discovering the environment.

Skinner (Burrhus Frederic Skinner was a well-known but controversial psychologist who founded behaviorism) in particular, claimed that a person’s behavior can be shaped and guided through such a method of controlling rewards and punishments. He did not worry much about which consequence was the stronger one as he believed that if a behavior is reinforced, it was apt to be repeated. However, he does favors and believes that a positive reinforcement was more effective than punishment.

Essentially, behaviorism concentrates on the study of overt behaviors that can be observed and measured [Good, 1990]. It views the mind as a ‘black box’ in the sense that response to stimulus can be observed quantitatively, totally ignoring the possibility of thought processes that occurred in the mind. In essence, three key assumptions underpin the behaviorism view of learning:

1. Observable behavior rather than internal thought processes are the focus of study. In particular, learning is manifested by a change in behavior.
2. The environment shapes the learner’s behavior. Learning depends on the elements of the environment, not the learner.
3. The principle of contiguity (how close in time two events must be for a bond to be formed) and reinforcement (any means of increasing the likelihood that an event will be repeated) are central to explaining the learning process [Merriam, 1991].

In behavioral learning theories, the transfer of learning is a result of generalization. Situations involving identical or similar features allow behaviors to transfer across common elements. The arrangement of stimuli and consequences within the environment form the most critical factor in learning. Hence, behaviorists usually attempt to prescribe

strategies that are most useful for building and strengthening stimulus-response associations [Winn, 1990] such as instructional cues, practice, and reinforcement.

Critiques on behavioral learning theories assert that the ethical consequences underlying behaviorism are great. Under the eyes of behaviorism, man is stripped of his responsibility, freedom, and dignity, and is reduced to a purely biological being, to be "shaped" by those who are able to use the tools of behaviorism effectively. It is also generally agreed that behavioral principles cannot adequately explain the acquisition of higher level skills or those that require a greater depth of processing (e.g., language development, problem solving, inference generating, critical thinking) [Schunk, 1991]. Furthermore, as behaviorism is stimulus – response based, the learning is strongly dependent on having and maintaining the appropriate stimuli to continue the intended behavior. Thus, if a certain incentive or stimuli is not present, then the expected behavior cannot materialize. For example, an assembly line worker, conditioned to react to certain stimuli, may not be trained to adapt to unforeseen situations.

2.1.8.2 Cognitivism

Cognitivism, as a learning theory, uses an information processing model to consider learning. It addresses the issues of how information is received, organized, stored, and retrieved by the mind. It also focuses on the complex cognitive processes such as thinking, problem solving, language, concept formation and information processing [Snelbecker, 1983]. Cognitive theories stress the acquisition of knowledge and internal mental structures [Bower, 1981]. Knowledge, as viewed in cognitive theories, is the symbolic, mental constructions in the minds of learners and learning becomes the process of committing these symbolic representations to memory where they may be processed. Hence, learning is concerned not so much with what learners *do* but with *what* they know and *how* they come to acquire it [Jonassen, 1991]. Learning can be either direct or implicit. Direct learning is intentional while implicit learning occurs when information sneaks into the learner's mind without the learners necessarily being aware of them, and without any explicit teaching. Different states of working memory are also emphasized. Consciousness and short term working memory stores the learner's current thoughts and reflections while persistent memory and knowledge is stored and organized for lifetime

use, although it can decay over time due to the lack of use. Often, learning involves revising the learner's mental schema to accommodate new, conflicting information.

Cognitive theorists view learners as active participants in the learning process who are concerned with the internal mental processes of the mind and with how it can be utilized in promoting effective learning. The main principle underlying cognitivism lies in the prominent role that the memory plays in the learning process. While cognitive theories also emphasize the role that environmental conditions play in facilitating learning, they contend that the environmental "cues" cannot account for all the learning that results. The most important element, they advocate, lies in the assimilation of new information into the learner's 'prior knowledge'.

Prior knowledge is a combination of the learner's pre-existing attitudes, experiences and knowledge. Such cognitive information (thoughts, beliefs, attitudes, and values) is considered to be influential in the learning process [Winne, 1985]. As such cognitive information is the belief of the learner, it can include misconceptions as well as knowledge that other people generally agree with.

In cognitive learning theories, the transfer of learning is a function of how information is stored in the memory [Schunk, 1991]. When a learner understands how to apply knowledge in different contexts, then transfer has occurred. As cognitive theories emphasize the relating of information to the learner's existing knowledge in memory, cognitivists usually attempt to prescribe strategies that organize information in such a manner that learners are able to connect new information with existing knowledge in some meaningful way. Cognitive learning theories are often employed to aid in the complex form of learning like reasoning, problem-solving and information processing.

A major weakness of cognitivism lies in its over-emphasis on prior knowledge. While prior knowledge aids meaningful learning, a learner, brought up in a cognitivism setting, will be at a disadvantage whenever relevant prerequisite knowledge does not exist. Another critique on cognitive learning theories asserts that cognitivism is similar to behaviorism in the belief that there are only finite, pre-determined goals. Having pre-

determined goals may be in fact desirable since it offers clear direction and purpose but such a fixed set of expectations can limit the potential of learning. Learners and instructors may become satisfied with obtaining minimum competencies or carry the attitude that “if it’s not broke, then don’t fix it!” when the learning experience could actually be designed better [McLeod, n. d.].

2.1.8.3 Constructivism

Constructivism, as a learning theory, is founded on the premise that learners learn by reflecting and constructing their own understanding of the world they live in. Instead of the objectivistic philosophical assumptions underlying both the behavioral and cognitive theories (i.e. the world is external to the mind and knowledge is reliably based on observed objects and events), constructivism views learning that starts off from the learner’s experiences. It emphasizes the ‘building’ that occurs in the learners’ minds when they learn.

Learning, as suggested, is an active rather than a passive process as learners construct their own unique mental model by combining information in their minds with the information that they receive from their sense organs. In essence, constructivism equates learning with creating meaning from experience [Bednar, 1991]. Each learner, as an individual, views the world in ways like no other learner and generates their own rules and mental models which they use to make sense of their experiences. Hence, learning is simply a process of adjusting their mental models to accommodate new experiences.

In constructivism learning theories, the transfer of learning must be facilitated by tasks anchored in meaningful context. The learning context becomes meaningful only when the context forms an inexorable link with the knowledge embedded in it [Bednar, 1991]. Also, as constructivism places great importance on experience, meaningful context depends on the authenticity of the experience. The authenticity of the experience hence becomes critical to the individual's ability to apply the learning [Brown, 1989].

The key aspects of constructivism are summarized by Woolfolk [Woolfolk, 2004]:

- Engage learners in complex learning environments and authentic tasks (situated learning)
- Social negotiations with experts and peers to help construct knowledge
- Multiple perspectives and representations offered (in different contexts and ways)
- Learners made conscious of their active role in learning
- Learners take charge of their own learning

Like cognitivism, a major weakness of constructivism lies in its over-emphasis on an individual learner's interpretation of learning content.

2.2 Research Work

From the experience of the past, we have learnt that e-learning is manifestly not a substitute for real teachers in real classrooms or the current irrational all-encompassing solution that many perceive. With its potential to make a huge impact on both academic and corporate learning, e-learning is fast becoming a proxy for any training associated with advancing performance, better return of investment, knowledge and skills impartation. With such an attractive alternative approach to traditional education, it is not difficult to comprehend why the educational community places such high hopes and efforts on e-learning. However, in an attempt to satisfy too many needs at a go, this research advocates that the concept of e-learning in the field of educational research and the learning sciences has become so broad in its meanings that it has become unclear in its significance. In theory, e-learning envisages the enhancement of traditional learning through the application of modern technology to provide high-quality, personalized, user-centric, easily accessible, on-demand, enhanced cognitive development, interactive learning and performance support. This is surely too much complexity to take on at once. Such massive demands have unconsciously placed too much expectation on the e-learning systems. We argue for an urgent need to redefine e-learning and to find a sound learning pedagogy for implementing e-learning solutions. It is only through this reflection, redefinition and research process that we can shift the focus away from the current irrational vision to one that is down-to-earth and achievable. This, we believe, will finally realize e-learning's true potential as an imperative supplement and enhancement of traditional methods.

Following this, our research direction is discussed. Evidence supporting our research direction is provided. This is followed by summarizing the key research areas pertaining to our research.

2.2.1 Research Direction

This research focuses on addressing the fundamental issues of learning. Given that e-learning starts out as a learning solution, we investigate how e-learning, relying on the advancement of technologies, has aided in the provision of learning. Specifically, we argue that the current e-learning solutions, in their attempt to answer too many needs at one time, have lost their focus. In the race for e-learning adoption, many organizations are minimizing or overlooking the importance of the basic principles of learning and focusing on other priorities. Gone are the considerations that e-learning, with or without the notion of ‘e’, is about learning. Priorities have certainly been misplaced. Here, we present concrete evidences of misplaced concerns.

2.2.1.1 Technology Concerns Precede Pedagogical Concerns

The increased need for learning and the belief that technology is a cornerstone in the effort to meet this need has convinced many organizations to pin their faith in e-learning. Using technology for the purposes of e-learning appears to be an attractive proposition. It is undisputed that the Internet and ICT has revolutionized learning. The ease of creating and delivering interactive, on-demand multimedia learning content has certainly enhanced learning. Different types of technology are supporting a wide range of educational settings and learning situations. Many organizations today think in terms of e-learning when implementing their training via either an e-learning portal or an LMS. LMS has taken the central role in providing a one-stop system that offers functionalities for both the content providers – i.e. administrating courses, statistics tracking, communication and collaborative learning possibilities – and the learners – i.e. digitalized and interactive learning contents, discussion forums and annotation facilities.

However, is learning all about technology? The dominance of technology over pedagogical issues has caused growing concern amongst pedagogists [Mehanna, 2004; Gibbs, 2006]. Pedagogists have been increasingly concerned about the widening gap

between educational theories and existing learning environments whose development they believe is driven mainly by technological advancements rather than educational objectives [Renger, 2002; Gibbs, 2006, pp. 48]. Also, much of the potential for e-learning is lost because too much of the pedagogy for e-learning has been transferred unreflectively from didactic traditional teaching where the computer substitutes for the teacher and textbook as the conveyor of information.

The ‘e-’ before the learning has misled many [Lytras, 2001; Chee, 2004] and e-learning seems to suffer from the heritage of the many ‘e-’ things that in the last decade formulate the new business environment [Lytras, 2002]. We advocate that e-learning should be more than a collection of technological solutions. Technology itself does not improve learning [Alexander, 1998]. Criticisms of e-learning systems a few years back by Rosenberg [Rosenberg, 2001] include poor content quality, unauthentic learning, emphasizing the form of learning over its substance, straitjacket instructions that do not cater to learners’ needs, boring learning content or ‘shovelware’ (that is, traditional paper-based material mounted directly on the Web in digital format). Also, no underlying pedagogical principles have been included [Bixler, 2000]. While current learning contents have improved in terms of its presentation, delivery and its interactive content, it is important to note that all its progress has occurred because of technological advancements. Noteworthy is the fact that the pedagogical status of current e-learning systems remains stagnant.

The current pedagogy of e-learning, if any, has adopted the same pedagogical principles that have been applied in the traditional learning context. Many organizations have put faith in the traditional pedagogical model that is characterized by activities such as lectures, tutorials and even laboratory work. Except for the fact that these learning materials have been digitalized and placed online, these pedagogical principles have not been extended to accommodate and provide for the dynamic nature of e-learning.

The current situation seems to be due to a war of sorts between technology and learning (pedagogy) and is to a large extent unavoidable. Most content developers are actually programmers who adopted the paper-based learning materials wholesale and transformed

them into digitized, interactive materials. These content developers think in terms of programming logic and the technical (delivery and networking) details. They are not skilled in the art of teaching nor are they concerned on the purpose of teaching and the philosophy to inform. The content experts on the other hand, while skilled in teaching pedagogies, are usually programming or even IT illiterate. Also, most content experts are experts only in the field of traditional training. Without the benefits of face-to-face communications, they are disabled and cannot effectively transfer their teaching expertise onto the digitalized learning materials. Furthermore, these two groups with their different expertise essentially speak different languages and their design and development styles are at odds by nature.

While most tools for e-learning such as LMS claim to provide the middle ground to bridge these gaps and to garner the benefits of both, findings have shown that most LMS vendors have deliberately distanced themselves from pedagogical issues [Firdiyeyek, 1999]. This is ironic. While most vendors claim to provide the complete e-learning solutions through every possible inclusion of technologically feasible features, there is an obvious absence of overt pedagogical integration. The main question to ask is: how can complete e-learning solutions exist without pedagogical techniques? At best, these LMS are mere providers of technology. What are lacking, and perhaps the most important, are guidelines on how to design, develop, deliver and manage pedagogically sound e-learning learning resources.

Some of the demands expressed have been fulfilled in recent generations of e-learning through the development of Learning Content Management System (LCMS). LCMS are designed with the intention of enabling the content experts, with little technological expertise, to design, create, deliver and measure the results of the e-learning courses rapidly. LCMS applications fundamentally change the value economics of e-learning content delivery by offering organizations a scalable platform to deliver proprietary knowledge to individual learners without bearing a prohibitive cost burden [Robbings, 2002]. However, while LCMS attaches substantial attention on the content and offers some form of pedagogy consideration, only one form of pedagogy is in use: that is, in order to learn, the learner has to work through a sequence of learning objects. This

pedagogy works under the underlying assumption that learning is a process of consuming content and teaching is envisioned as the art of selecting and offering content in a structured, sequenced way and the learner's progress can be gauged just by tracking his progress through the sequence of learning objects [Koper, 2004]. However, current educational practice is more complex and advanced than this. Besides the multi-interaction involved, pedagogical approaches that incorporate social/affective, metacognitive, cognitive and constructivist strategies must be considered.

To this end, we advocate that, technology, no matter how advanced, should be integrated into learning. Rather than perceived purely as a technical aid to provide a new flexible delivery medium, current 'technology-integrated learning' must address the very important issues of learning pedagogy to enhance the learning and the teaching process.

2.2.1.2 Instructor's Needs Precede Learner's Needs

The traditional pedagogic model is characterized by teachers who assume the full responsibility for making decisions about what will be learned, how will it be learned, and when it will be learned. The teacher directs learning. However, noteworthy is the fact that great teachers of ancient times, from Confucius to Plato, did not pursue such authoritarian teaching techniques. How such teacher-focused learning later came to dominate formal education is amazing.

"Our academic system has grown in a reverse order. Subjects and teachers constitute the starting point, [learners] are secondary. In convention education the [learner] is required to adjust himself to an established curriculum... too much of learning consists of vicarious substitution of someone else's experience and knowledge" [Conner, 2003]

As a result of such authoritarian teaching techniques, it is very common, as Knowles [Knowles, 1970], way back in the 1970s, has pointed out:

"Student to leave school adult in other ways, but still dependent, or at least retarded in independence, as a learner" [Knowles, 1970]

Unfortunately, as we step into the 21st century, the situation has not improved. In view of the negative aspects of such teaching techniques, the inception of e-learning has seemingly called for the abolishment of such teaching methods by exploiting the advancement of technologies to incorporate both the teacher's and the learner's perspectives. While teachers or content experts can still produce learning materials from scratch as per the traditional learning model, they can also search, reuse or repurpose other available learning resources to cater to their own teaching objectives. The ability to reuse is one main aspect that makes e-learning attractive as it not only minimizes development time but also provides a better ROI. More importantly, this reusability aspect of e-learning frees up the teachers' time and allows them to concentrate on providing personalized content and support for the students.

In fact, we feel that it is this ability of e-learning to provide a learner-centric and personalized learning environment where the learners, for the first time, take center stage and dictate their own learning that makes it so attractive. Perhaps, it is the very nature of e-learning (i.e. minimized teacher-student interaction, geographical separation between teacher and student) that calls for such form of learner-centric learning. Nonetheless, such learner-centric style of learning is what we truly believe to be consistent with what we know to be essential for our prevailing educational philosophy.

Unfortunately, though the providing of learner-centric learning resources is essential for e-learning, and for the fundamental process of learning in general, it is difficult. While most current e-learning systems and providers proclaim learner-centric, personalized, and vibrant learning environment created through blending different appropriate teaching strategies and technologies, most e-learning content still focuses on the content experts and not the learners [Angehrn, 2001; Abramowicz, 2003]. Current courses, although online and automated, are still delivered within the boundaries established by the subject matter experts (SMEs) who play the role of both the director and dictator of the learning experience. Most e-learning systems are focused on what the SMEs offer. The learning paths are still rigid and restricted. Minimum or no learner's preferences are being taken into account. For example, once the learner is enrolled into an e-learning course, the learning will become passive and all the learner has to do is to simply follow the

prescribed paths through the whole courseware (dictated once again by SMEs) right from the pre-assessment until the post-assessment. This style of learning brings us back to the days of traditional classroom learning except for the fact that learning now takes place online instead of being restricted by the place and time. The tendency to simply replicate traditional structures of classroom teaching and to make the instructions online adheres to what Barr and Tagg [Barr, 1995] refer to as the Instruction Paradigm. While the Instruction Paradigm (or the SME-centric approach) may be appropriate for the traditional classroom learning, the “distance” component in e-learning makes such approach to instruction less effective [Chee, 2004]. Effective learning must produce deep understanding, not merely knowledge reproduction. Hence, with most e-learning programs organized around the needs of SMEs, the learners are left to tolerate the homogenized, standardized subject matter.

2.2.1.3 Distribution Concerns Precede Learning Concerns

While current e-learning system claims to provide knowledge acquisition, storage, retrieval and maintenance through the uses of metadata and standards, this research advocates that most e-learning systems at best, serve only as a form of information repository. Current status of e-learning systems is more of information access and distribution more than e-learning. For information to transform into knowledge and be synthesized into ones cognitive structure such that it becomes ‘operational’ and deployable requires deep internalization and learning. Current e-learning systems and courseware has failed miserably as they are often monolithic, inert and fail to facilitate the development and sharing of knowledge [KnowNet, 2004]. Simply enabling ready access to pertinent information, or even instructional information, need not necessarily entail learning or education [Chee, 2004].

“If access is all that matters, education would be reduced to a trivial matter of simply providing an excellent library then all the learners will get an educated citizenry” [Chee, 2004]

Such emphasis on distribution rather than learning highlights the misplaced concerns. Instead of considering how to impart learning, many content providers now think in terms of ROI when developing their contents.

2.2.2 Key Research Issues

The evidence we presented here is clear: new technologies, however effective in other fields, don't inevitably lead to major change or improvement in education. While the type learning environment affects the style of content presentation and teaching, it does not affect the fundamental processes of human learning. Regardless of the type of learning environment, the most important process is to enable the students to actively engage with the learning materials and to be able to synthesis the new information into the student's cognitive structure. Hence, it is important to note that the real issue of e-learning is about moving from a teacher-centered environment to a learner-centered environment; NOT from traditional to e-learning. Table 2.1 summarizes our findings and presents the gap between the ideal e-learning methodology and the current practice.

Ideal E-learning Methodology	Current Practice	Assessment of Current Practice
Technology: Leverage on the advancement of communication and information technology to deliver and enhance learning and teaching	Employed technology to provide <ol style="list-style-type: none"> Media-rich, interactive, personalized content On-demand learning that transcends geographical and time boundaries 	Technology has enhanced learning and teaching by overcoming the learning/teaching limitations that were traditionally deferred due to technology constraints.
Pedagogy: Employ the proven learning theories (of the past) and pedagogies to provide strong structured foundations for planning and conducting instructional design activities	The current pedagogy of e-learning, if any, has adopted the same pedagogical principles that have been applied in the traditional learning context and is characterized by activities such as lectures, tutorials and even laboratory work.	Much of the potential for e-learning is lost because too much of the pedagogy for e-learning has been transferred unreflectively from didactic traditional teaching where the computer substitutes for the teacher and textbook as the conveyor of information. [Phillips, 2005; Mason, 2006; Gibbs, 2006]
Personalized Learning: Provide a learner-centric and personalized learning environment through blending different appropriate teaching strategies and technologies where the learners dictate their own learning	Current e-learning courses, although online and automated, are delivered within the boundaries established by the SMEs who play the role of both the director and dictator of the learning experience. Most e-learning systems are focused on what the SMEs offer. The learning paths are rigid and restricted. Minimum or no learner's preferences are being taken into account.	Most e-learning systems and providers cannot provide learner-centric style of learning and still focus on the content experts and not the learners [Angehrn, 2001; Abramowicz, 2003].
Reusability: Employ learning object philosophy where content providers moved from creating and delivering large inflexible training courses to database	Actively championed by Cisco Systems and supported by IEEE LOM – through the use of granularity or aggregation level to define and determine reusability	E-learning is centered on the issue of reusability and the concept of RLO. However, the physical implementation of the RLO concept is difficult. Today,

driven objects that can be reused, searched, and modified independent of their delivery media		there are no known e-learning systems that are completely reusable.
---	--	---

Table 2.1 Gap between the Ideal E-learning Methodology and Current Practice

Our findings have shown that current practice has only succeeded in eliminating the limitations of time and space for learning transactions. Gone are the motivational aspects inspired by the teachers' physical presence and their ability to respond to the cues presented by live audiences. Hence, despite e-learning having the potential to reconceptualize what seems to be impossible in learning and teaching, this enormous potential is not realized fully.

Amidst the many aspects of e-learning that need further studies and improvements, this research focuses only on the **learning personalization** aspect. Specifically, the following key issues are addressed:

1. The dominance of the teacher-centered approach needs to be challenged if the best of what technology offers is to be realized;
2. The alignment of all stakeholders in the design and delivery of courses must be achieved – not only in terms of ROI but in terms of organizational learning goals;
3. The over-simplistic view that any content, packaged and delivered on the Internet, constitutes e-learning learning resources;
4. The promise of personalized delivery of learning content to match learners' learning styles, needs and goals must be fulfilled;
5. There is a vital role for pedagogy (and scaffolding) to play. However, it is in danger of being neglected in the rush to make all things possible and fast in the open environment of the Internet.

It is important to note that the learning that results from e-learning depends much more on the skills of the content experts rather than, as is often implied, on the number of sophisticated features present. Learning need not be driven by technologies or even new pedagogies. What is important is to develop novel forms of educational processes and knowledge structures while carefully maintaining and enhancing the pedagogical principles that remain fundamental to almost all forms of learning. Hence, it is not

important to have new pedagogies but to utilize and adapt previous pedagogies effectively and efficiency to existing e-learning needs. Also important is the role of learning technology. Advances in digital media have made this an exciting business but we must not fall into the trap thinking that multimedia presentations, even those which permit the learners to immerse themselves in virtual worlds, are actually teaching. What is being learned from any presentation of information depends far more on what the learner already knows than on the extrinsic properties of the presentation.

2.3 Review of Related Work

2.3.1 Importance of Learning Personalization

Educational research informs us that ‘one size does not fit all’ [Reigeluth, 1996]. Learning characteristics of students differ [Honey, 1986] and different students, reflecting on their own traits, preferences, culture and upbringing, process and represent knowledge in different ways [Riding, 1998]. They also have their preference for different types of resources and exhibit consistent observable patterns of behavior when they learn [Riding, 1998]. Research has also shown that it is possible to diagnose a student’s learning traits and that some students learn more efficiently when instructions are adapted to the way they learn [Rasmussen, 1998].

However, our current utilization of technology and contemporary pedagogical principles are not adequately addressing the shortcomings of our educational sector. Even with our exploitation of technology for learning, there still exists a significant gap between what the learners need and what we can currently provide. Besides, our era is characterized by knowledge as the critical resource, constant technological change and time constraints. Hence, 21st century students demand the flexibility and security of taking responsibility over their learning. Gone are the days where educators should dictate what, how, and when to learn. Students are taking pride in determining and controlling their unique learning roadmap and can no longer accept homogenous learning materials nor leave their learning completely in the hands of others.

Notwithstanding that the purpose of learning is to provide opportunities for learners to think critically, the learners of today and tomorrow will face an information-rich

environment in which new expectations are demanded. Our research has shown that at the heart of these expectations lies the personalization of learning experiences. Personalizing instructions and interactions to address an individual learner's needs is one fundamental cornerstone of today's knowledge-centered learning paradigms and has received considerable attention [Mobasher, 2000; Herlocker, 2001; Rashid, 2002].

In fact, learning personalization when placed in the context of e-learning has to take on an even more important role in assisting the act of learning. Besides having to fill the gap left by the reduction or elimination of face-to-face interactions, learning personalization also has to take over the responsibility of the teacher. While the learners take pride in dictating the pace, amount of content and setting the place and time of learning, they are currently not equipped with the skills to assume full responsibility over their own learning. Neither are they equipped with the mindset to effectively plan their learning paths. Hence, they often experience a lack of the important learning factors of motivation and persistence [Súilleabhain, 2004] and have expressed feelings of isolation, lack of self-direction and increased management problems [Harasim, 1995; Abrami, 1996; Bennett, 1999] when they are called upon to take control over their learning. Furthermore, in order for them to succeed, they are assumed to possess some form of self-discipline, ability to work alone, good time management, learning independence, readiness, the ability to plan for themselves and the ability to assess their strengths and weaknesses [Dunlap, 2003; Watkins, 2005]. This is surely too much for the learners to take on at once.

Such an ironic situation is perhaps understandable. Most learners are brought up in a traditional education approach where the learners' learning activities, curriculum and learning routes are largely static, constrained and prescribed by the teachers. Little, if any, opportunities exist for the learners to plan their own educational goals or to consider their own learning priorities, needs and preferences. Often, in a traditional setting, the learners are spoon-fed and the exact topics to be learned are made explicit. However, for e-learning, such guidance is reduced. Instead, the learners are encouraged to construct their own learning paths, consider their own needs and learning goals and to rely less on the teacher for direction. This type of transition from a teacher-centered approach to a learner-centered approach is applauded by many. However, while this type of transition is

essential, the evolution is too abrupt. The sudden influx of freedom coupled with a lack of guidance and support has made the students feel daunted. As Hammond and Collins [Hammond, 1991] points out, “learners accustomed to teacher directed learning may have no experience of self-management of learning so it may initially be intimidating.” Furthermore, the skills (the ability to plan, the ability to manage learning, the ability to review and reflect on reasoning and research skills, etc.) that are often associated with self-centered learning have placed too much cognitive demands on the students. Another issue, given such freedom, is the lack of awareness of the amount of knowledge or even the specific syllabus to follow.

Problems with regard to stimulating and sustaining learner motivation are also well documented in the literature of e-learning and the narrower context of distance learning [Zvacek, 1991; Rowntree, 1992; Visser, 1998], especially when learners are working independently at a distance. The learner’s motivation is important for the learning process as it has been shown that students will only restructure or assimilate new data only if accommodation fails and when he or she is motivated to reconcile anomalies and to reduce inconsistencies [Wankat, 1993]. However, overcoming these motivational challenges can be difficult because of the complexity of human motivation and the vast number of motivational concepts and theories that exist [Keller, 2004]. While the challenge of motivation and drop-out rates are typically answered through the provision of traditional face-to-face communications [Súilleabhain, 2004], the luxury of such provision is not available, or at least minimized, in the context of e-learning.

Hence, in view of the preceding discussions, we argue that learning in an open environment such as e-learning demands a more effective and personalized approach to provide learner orientation and individualized access support. Such support is crucial. It is advocated that such personalized learning support will have a positive effect on the learners in terms of motivation, retention, self-direction and learning performance. Research has shown that personalization is one of the key factors that is directly related to user satisfaction [Riecken, 2000] and has been shown to be useful in several areas such as e-commerce [Kasanoff, 2001], business-to-business companies [Colkin, 2001] and e-learning [Mor, 2004]. However, it must be introduced carefully [Nielsen, 1998] as the

success or failure of the performed activity hinges heavily on the factor of personalization. Table 2.2 summarizes the key characteristics of personalized e-learning system and their merits.

Characteristics	Merits
Learning Preference Parameters Considerations i. Learning Styles – i.e. visual/verbal ii. Learning Expertise – i.e. beginner, expert level iii. Learning Constraints – i.e. course fees, duration	Personalization seeks to 1) understand the deep-seated psychological impacting differences in learning behavior, 2) make predictions about delivering content, 3) deliver content specifically to help the learner achieve learning objectives and more importantly, 4) improve learning ability and enhance online learning relationships. Cognitive-based personalization system uses information about learning preferences or styles from a primarily cognitive perspective to deliver content specifically targeted to differing learner attributes. This type of personalization operates on more complex algorithms than the other types but is able to factor more learner attributes and enhance the learning experience.
Assisting the act of learning and learning path / curriculum planning	Personalized curriculum sequencing is an important research issue for Web-based learning systems because no fixed learning paths will be appropriate for all learners [Hong, 2007]. Therefore, it is important for the personalized e-learning system to assist in the act of learning and learning path/curriculum planning and take the guesswork out of learning. By adopting a straightforward yet pedagogically-rooted methodology, personal learning visual maps are recommended to integrate a variety of learning aids to optimize learning.
Scaffolding Support	The gap left by the reduction or elimination of face-to-face interactions in e-learning system is evident. To address this and to ensure meaningful interactivity, there is a need to find adequate and appropriate guidance and scaffolding support. This is important to help learners control and respond to the content while actively making and constructing their own meaning from the content [Tang, 2005].
Goal-oriented learning that stimulates and sustains learner motivation	e-learning does not follow the traditional schema of goal setting and classroom activities. Here, the 'teacher' is not giving clear and concrete goals for learning tasks, but only providing context and starting point for students' inquiry. As such, the learners are required to show more self-regulative and are responsible of generating their learning agenda and setting up their learning goals. Further, learning tasks are more open and ill-structured and thus more complicated to work with. Therefore, a personalized system must include socio-cognitive and socio-emotional goal restructuring. That is, the system must focus on how goals transform into action – dynamic interplay of personal beliefs,

	situational interpretations and subsequent actions – so as to improve, stimulate and sustain learner motivation.
Pedagogical Concerns	Curriculum re-designs to support learning personalization in learning and teachings go beyond setting up the system infrastructure. Learning is rooted heavily in pedagogical paradigm. As such, only pedagogy can make two diverse disciplines of education and technology synergize in the deployment of technologies in teaching and learning.

Table 2.2 Key Characteristics of Personalized E-learning System and their merits

Envisioning the shift in future e-learning direction from a content-oriented approach to a personalized learner-centric and knowledge synthesis approach, we believe that effective knowledge personalization, synthesis and its transfer is the key competitive differentiator that governs an e-learning system.

2.3.2 Learning Personalization Status

Successful implementation of personalized e-learning systems is thin on the ground. However, learning personalization is not a new learning notion. Research work on learning personalization started as early as 1962 [Glaser, 1962; Mager, 1962]. Typically, early efforts emphasize performance objectives and deliberately use events to constituent external conditions of learning so as to stimulate internal learning processes. Also it was around this time that there was a shift in focus towards the use of media as part of the educational process instead of as the primary focus of instruction. However, during that period, learning personalization was still both technologically and physically impossible. Thus, instructional design theorists only focused on the principles of their learning models and did not make any suggestions for flexibility and personalization of a learning experience that is based on the learners' context. Hence, in spite of the vast efforts spent, these efforts contributed only to the field of instructional systems.

It was only until the emergence of online or Web-based learning environments that research on learning personalization took on a greater emphasis. The increased efforts being placed on research in learner-centric, web-based instructions has prompted many researchers to begin examining the effects of personalization in an educational setting [Reiser, 1987; Boud, 1992; Khan, 1997; Brusilovsky, 1999; Martinez, 2000; Chou, 2000;

Lee, 2001; Kao, 2001; Lin, 2001; Papanikolaou, 2002; Henze, 2005; McLaren, 2006; Styliadis, 2006; Turker, 2006]. Theoretically, while current advancements in informational technologies offer opportunities for precision identification of learners' learning needs, expertise, highly personalized content and individualized support, much still needs to be learned both technologically and pedagogically.

Existing e-learning systems are still very much in the experimental stage of creating a personalized learning environment. While personalization is claimed by almost all e-learning systems, personalization is only realized superficially. Currently, the amount of personalization does not go beyond the simple customized greetings (with the learner name) or some suggestions of newly added courses. At best, these systems incorporate some form of simple agent that recommend some courses based on the learner's previous searches and selections. Although some learners might appreciate these personalized features, these small enhancements do not necessarily enhance their learning experiences.

Also, most e-learning systems employ conventional and intuitive methods of course development and do not provide learners with the locus of control in their learning. Hence, without the ability to capture learners' needs at real-time, these systems simply produce a one size fit all course and leave the learners to tolerate the homogenized, standardized subject matter as highlighted in chapter one. Such presentation methods, at best, can only facilitate rote learning [Novak, 1984] but not meaningful learning [Ausubel, 1965].

Besides, the open problems of supporting learners' identification and profiles dynamically, in a distributed environment, remain unsolved. Hence, despite the potential of information technology to provide optimum delivery of dynamic personalized learners' interactions, existing systems fully dictate the pace of instructions and offer no option for a learner's interaction at real-time.

2.3.2.1 Current Learning Personalization Scenario

As an illustration of the current limitations in learning personalization efforts, a typical learning scenario set in a Singapore context is presented. This scenario is employed to motivate the thesis with a concrete application. Subsequently, this scenario will be

revisited throughout this thesis to illustrate the theories, algorithms and methodologies that are presented in the later chapters. This serves as concrete examples that come directly from the application of our proposed solutions.

The scenario is based on the Singapore Educational System and follows the Ministry of Education learning guidelines (see <http://www.moe.gov.sg/>). A Primary 6 student encounters a mathematics problem which requires him to compute the ratio between the areas of two triangles. He surfs the Internet in hope of finding courses that impart knowledge to solve the problem. He manages to find separate course offerings that teach ratio and area. However, the courses provided do not cover all the learning concepts involved. For example, the discussion on ratio relies heavily on the concept of percentage and proportion which he has limited knowledge on. Frustrated, he moves on to the course on area but also encounters the same problem as that course assumes pre-requisite knowledge on the topic of geometry. This scenario depicts a typical example of what our education system is offering. The current comprehensive learning systems using their ‘proclaimed’ personalized search engine are at best, able to find the appropriate course offering. They cannot tailor the course according to the learner’s expertise. Any form of personalization does not extend beyond the information gathered from the learner profile; i.e. learner name, past courses attended, etc.

Extending this scenario in the corporate context, a personalized e-learning system must take on a more holistic role to learning customization. For example, when there is a new project, there exists a need for the e-learning system to automatically suggest a project team, measure the human resources and competence gaps and then reduce the competence gap by creating personalized learning paths to equip the team members with the relevant project skills. Moreover, the system should be able to dynamically redefine the proposed learning paths according to feedback in order to optimize the acquisition of the needed competencies. However, current systems again suffers from the abovementioned problem; that is, they cannot tailor the course according to the learner’s expertise. In the corporate case, current systems cannot effectively plan for learning/teaching. Often, these classical planning systems are not transparent as the

details about learning plan generation are hidden inside software components. This makes it difficult for pedagogists to understand, and in consequence, to trust and use them.

2.3.2.2 Learning Problems Identified

From the scenario, the following learning problems are highlighted.

1. A one-size-fits-all learning solution for all learners
2. Recommended solution does not take the learner's learning preferences or expertise into account
3. Inaccurate learning solution recommendation due to keyword match technique
4. No learning support provision
5. Stand alone learning solution with no follow-up recommendation to other relevant or related courses
6. Teaching plans are not transparent

2.3.3 Overview of Existing Learning Personalization Approaches

In order to address these learning problems, most personalized systems consider learner preferences, interests, and browsing behaviors when providing personalized services [Chen, 2005]. Existing systems can be classified into three categories: Search Engines, Recommender Systems, and Adaptive Hypermedia Systems.

2.3.3.1 Search Engines

The rapid growth of the Web [Lawrence, 1998] led to the problem of information overload [Berghel, 1997; Borchers, 1998] which caused many learners to waste precious time in navigating the enormous information space. Using the Internet to search for information has become a chore as learners have to sift through thousands of articles to find that only a few of them are relevant. Often, learners complain about spending too much time navigating the Web and yet being unable to find the information they require [Lawrence, 1998; Kobayashi, 2000; Arasu, 2001]. In fact, research has shown that many learners give up their search after the first try, examining no more than ten documents, which is what most search engine output on the first page [Jansen, 2000]. As a result, many Web applications are developing powerful search tools [e.g Brin, 1998; Chidlovskii, 2000; Kleinberg, 1998;] and engines [e.g. Google search engine, Yahoo,

AltaVista and CiteSeer Website] that allow learners to search more efficiently. Most of these tools and search engines provide some form of personalized mechanisms to enable learners to filter out uninteresting or irrelevant search results.

One of the personalized mechanisms currently employed for search engines is the request for contextual information about the search query. (see for example, Inquirus 2 project at NEC Research Institute [Eric, 2000], Outride System [Pitkow, 2002]). Often, a layer of explicit contextual information in the form of category restriction is added above the regular search engines. Examples of category restriction are educational background, domain of interest, previous user interaction, or classification questions such as ‘is the user looking for a company that sells a given product or is he looking for the technical specification of a given product’.

Such a category may considerably clarify a query and thus focus the search. Therefore, instead of the normal searches that treat each search request in isolation and present results that are identical, independent of the learner or the context in which the learner made the request, additional contextual information are used. For example, a normal search query of the term ‘Jordan’ will usually return results that are related to the famous basketball star, be it his shoes (Nike), his fan club, or his basketball history. However, when the learner is able to specify the context whereby the search query should reside, then the search becomes personal and more likely to cater to his needs. For instance, searching for the term ‘Jordan’ in the context of ‘academic’ and ‘homepage’ will probably return results of the homepage of the students or professors by the name of ‘Jordan’ and not the basketball star. There is also a family of tools that interpret the notion of context as a set of previous information requests originated by a user. Defined this way, context search becomes personalized, and tools in this category keep track of a user’s previous queries and/or documents viewed [Finkelstein, 2002].

Another common approach to personalized mechanisms is the usage of flat-ranked lists of query results [Ferragina, 2005]. It uses personalized ranking, an extension of the classical link-based ranking. Personalized ranking focuses on learner rather than the query by combining web queries with contextual/profiled information. Examples of such

personalized services include Google, Yahoo and Eurekster. These search engines collect category based profiles explicitly maintained by the learners. Patents have also been filed for such personalization services. For example, Google files a “Google Advertising Patents for Behavioral Targeting, Personalization and Profiling” patent on 6th October’ 2005 which looks at personalizing search results based on user profile. The user profile describes interests of the user and can be derived from a variety of sources, including prior search queries, prior search results, expressed interests, demographic, geographic, psychographic, and activity information.

2.3.3.2 Recommender Systems

Personalized recommender systems were first proposed and applied in the E-commerce area for product purchase [Lu, 2004]. Such systems help customers in finding suitable products through a list of recommended products [Cheung, 2003]. They are gaining so much popularity that recommender systems are changing from novelties used by a few E-commerce sites, to serious business tools that are re-shaping the world of E-commerce [Schafer, 1999]. While recommender systems have obtained success within the domain of E-commerce [Schafer, 2001], it has not gained the same currency in e-learning [Lu, 2004]. A recommender system in an e-learning context is a software agent that tries to intelligently recommend actions to a learner based on selected rules. In general, the rules can be classified broadly into either private-based or collaborative-based types. For private-based rules, the recommender will only recommend learning tasks based on learning tasks that have already been done by the learner (e.g. recommend the course of Java II after the learner has completed the course on Java I) or inputs that have been explicitly provided by the learner (e.g. an interest in programming languages may lead to the recommendation of the course on JAVA programming). Except from the learner’s interaction/inputs with the system, no other inputs will be considered. Collaborative-based rules are usually established by forming groups of learners. The similarity of learners could be established by using user profiles, or could be based on some common previous access patterns [Zaine, 2002]. For collaborative-based rules, learning tasks will be recommended through a common access model that recommends the same set of learning tasks to all members of the group.

Recommender systems can be implemented using a variety of techniques such as data clustering, association rule mining, intelligent software agents, collaborative filtering, etc. Some examples of exiting recommender systems include [Balabanovic, 1997; Lee, 2001; Fu, 2000; Rashid, 2002]. Typically, learner's preferences, interests and browsing behaviors are considered when recommending personalized services.

Generally, recommender systems in an e-learning context make individualized recommendations of the reading materials for the learners. The reading materials are usually collected and kept in a dynamically evolving paper repository. The basic idea behind the recommender is simple and is best summed up in Basu et al. [Basu, 2001], "*Given a representation of my interests, find me relevant papers*". Most recommender systems [e.g. Bollacker, 1999; Woodruff, 2000; Basu, 2001; McNee, 2002] use the citation matrix to make recommendations and imbed a web crawler to maintain the repository (update new papers and remove older papers). Some recommender systems also adopt collaborative filtering techniques [e.g. Konstan, 1997; Yan, 1999; McNee, 2002] while work has also been done to investigate methods that are able to recommend pedagogically suitable but yet poorly cited papers [Tang, 2003].

2.3.3.3 Adaptive Hypermedia Systems (AHS)

Adaptive hypermedia is a research area that stands at the crossroads of hypermedia and user modeling. A hypermedia application offers its users the freedom to navigate through a larger hyperspace. However, such a navigational freedom is daunting as users often get lost in the cyberspace. Hence, in order to offer a good compromise between offering navigational freedom and guidance support, adaptive hypermedia systems are investigated [Wu, 2001]. Adaptive hypermedia systems offer (automatically generated) personalized content and navigation support, so the choice between freedom and guidance can be made on an individual basis.

Adaptive hypermedia systems build a model of goals, preferences and knowledge of individual user (such as hyperspace experience, interests and individual traits) and use this model throughout its interaction with the user to adapt to the needs of the user [Brusilovsky, 2001]. It aims to couple hypermedia technologies with personalization

strategies to deliver course offerings that are tailored to each individual learner's learning requirements and styles [Conlan, 2002b]. AHS is also seen as the solution to the traditional 'one-size-fits-all' approach [Brusilovsky, 2001].

Through the separation of different forms of user and content models, AHS adapts the content and provides linkage of hypermedia pages to the user, thereby providing user personalization. For example, a student in AHS can be presented with information that is adapted specifically to his knowledge of the subject [De Bra, 1998]. A suggested set of relevant links can also be provided for the student to proceed further [Brusilovsky, 1998].

The user model can collect data about the user from various sources. The data collection methods can either be implicit – observing the user interaction - or explicit – requesting direct input from the user. Using data from the user model and intelligent adaptive technologies, AHS automatically tailors different interaction for different users even if the context is the same. AHSs are essentially collections of connected information items that allow a user to navigate from one item to another and search for relevant items. The adaptation effect in this reasonably rigid context is limited to three major adaptation technologies – adaptive content selection, adaptive navigation support, and adaptive presentation. When the user searches for relevant information, the system can adaptively select and prioritize the most relevant items. When the user navigates from one item to another, the system can manipulate the links (for example, hide, sort, annotate) to provide adaptive navigation support so that the users can easily choose where to go next. When the user gets to a particular page, the system can present its content adaptively [Brusilovsky, 2002].

2.3.4 Review of Existing Learning Personalization Approaches

The three approaches to learning personalization neglect the importance of learner/user learning ability. Learning personalization must take into account the different levels of learner/user knowledge, his prior learning experience in that particular context and also his learning preferences, traits, constraints and goals. All these considerations must be taken into account as they are related to the nature of learning [Brusilovsky, 1999; Lin, 2001]. As stated by Chen et al. [Chen, 2005], only when we are able to consider learning

ability, then we can promote personalized learning performance. Unfortunately, our existing methods fail to satisfy this requirement.

We further argue that most approaches address the issue of personalization, and not learning personalization. For example, the approaches used for search engines are more applicable to information retrieval systems rather than e-learning. These systems' main priority is to search the World Wide Web for information. No form of learning is taken into consideration. Personalizing information is often seen as presenting the best match between the search query and the retrieved information. The usage of contextual information seems to be another keyword that is added to the search and can hardly be seen as a personalization approach. Although some systems have worked around that problem by using the context information to select the search engines that process the query (e.g. Inquirus 2 project [Eric, 2000]), this process is difficult and limited as specialized search engines have to be created for each specific domain [Lawrence, 2000]. Furthermore, with thousands of specialized search engines, how can the system locate the search engine that is of interest to a particular user? More importantly, how much overhead or effort is required to locate the best specialized search engines?

The personalized ranking that is being adopted by many who claim to provide learning personalization only offers a partial solution. Personalized ranking only allows profiles over a tiny set of choices (i.e. Google and Yahoo). Moreover, it is resource-intensive as it needs to maintain up-to-date profiles which are a critical and private resource. In scientific literature, the personalized ranking problem has been investigated and scaling techniques have been proposed as solutions to the classical approach. However, such an approach does not solve the resource-intensive problem as the solution ultimately still needs to compute, for each Web page, a number of ranking values that are related to the number of user profiles.

The recommender system techniques although closer to the context of e-learning are unfortunately not very accurate [Zaine, 2002]. While they try to work around the resource-intensiveness problem by forming user groups, we argue that the usage of group profile defeats the purpose of learning personalization. Group profile, although catering

to a smaller group of learners, is not individualized anymore. Besides, we argue that it is difficult, if not impossible, to find perfect matching of learning preferences or goals between learners. Prior knowledge is one good example. No two learners have the same prior knowledge even if they are brought up in the same academic setting. Moreover, the initial classification of new user is still an unsolved problem in the domain of recommender systems [Rashid, 2002]. Furthermore, the recommender systems, while set in the context of learning, is not based on any proven learning theories. Neither are there any considerations on whether the learner can comprehend the reading materials that are recommended to them.

AHS, while set in a stronger pedagogical context, also suffers from numerous problems. The narrative or pedagogical model employed is usually either embedded in the content or inside the adaptive engine itself. Hence applying new or different pedagogical models, e.g. case based learning, simulations, etc., to the content model is more difficult and involves a re-authoring of the content model. This often results in learning content that is difficult to reuse or an engine that is domain specific [Conlan, 2002b]. Also, AHS is dependent on the successful capturing of individual traits such as personality, cognitive and learning styles. However, the capturing of such traits is difficult. Moreover, even if these traits can be accurately elicited, challenges on how to exploit this information still exists [Brusilovsky, 2001]. Specifically, key questions such as “how the systems can build and diagnose the student’s learning characteristics” or “how can they adapt the learning environment to suit the student’s needs” are still unanswered.

Also, most personalized systems collect information about the user implicitly. Often the users are not aware of it. Hence, this often leads to privacy problems [Kobsa, 2002]. Furthermore, most personalized systems use some form of predictive reasoning to personalize information. While such techniques can intelligently make recommendations, they are prone to error and a significant amount of time is required to fine-tune the reasoning rules for each learner before they can accurately prescribe personalized information.

Besides these three common approaches to personalization, recent effort started to incorporate new and more appropriate parameters for learning personalization that includes learner preferences, interests, activities and browsing behaviors [Huang, 2007; Conlan, 2007]. These parameters shift the focus onto the learner by considering feedbacks to learning styles [Vasilyeva, 2007] and learners' activities [Conlan, 2007] to provide better learning customization. However, these approaches still fail to integrate the context of learning that is defined by the pedagogical approach [Arapi, 2007] and neglect consideration on how learner ability and the difficulty level of the recommended curriculums are matched to each other [Huang, 2007].

The recent shift to Web 2.0 also has its counterparts in both e-learning technology and methodology – termed E-learning 2.0. E-learning 2.0 (coined by Stephen Downes) employs a 'small pieces, loosely joined' approach that combines the use of discrete but complementary tools and web services - such as blogs, wikis, and other social software - to support the creation of ad-hoc learning communities. These tools although not designed specifically for e-learning, can be used to empower students to create new and exciting learning opportunities [O'Hear, 2006].

Again, learning personalization is termed by the experts as central to learning in the future [Graham, 2007; NUS, 2007]. E-learning 2.0's approach to personalization is oriented towards the social networking principles and aim to harness the end user to add value to personalization. However, as this trend is relatively new, it currently only covers mass personalization that allow users with specific needs to access learning products, services or like-minded people [Rosen, 2006].

2.4 Summary

The era in which computers were first used for Computer Aided Learning and then E-learning (as an electronic form of information repository and presentation) is coming to an end. We have learnt that e-learning is manifestly not a substitute for real teachers in real classrooms. Rather, e-learning, we believe, will realize its true potential as an imperative supplement and enhancement of traditional methods, when teachers and learners begin to think seriously about its proper role in the pedagogical world. We

further argue that central to such role is the act of learning personalization that this research is focused on. Although current development efforts do not sufficiently address the important issues of learning personalization, emerging technologies and research efforts seem to offer some hope for resolving the key impediments to its successful implementation. Hence, we hope that by providing learning personalization, we can shift the focus away from the current unrealistic, all embracing one, to one that is down-to-earth, realistic and focused on objectives. We believe that e-learning has enormous potential for benefiting mankind, particularly those who are less fortunate and unable to afford an expensive education. The wide reach and potential for sharing that is possible through the Internet as well as the learning benefits that learning personalization brings should be actively exploited for this purpose.

CHAPTER 3

LEARNING PERSONALIZATION – A PEDAGOGICAL CONTENT DEVELOPMENT METHODOLOGY

This chapter will describe our learning personalization work from a pedagogist's perspective. This research work is done in collaboration with Professor Agnes Chang Shook Cheong, Associate Dean of the PGDE programme (Professional Graduate Diploma in Education) and a pedagogy expert in the National Institute of Education, Singapore. The chapter starts off by stating our stand on e-learning pedagogy. Specifically, the research gaps and our research contributions to the e-learning pedagogy field are stated. Section 3.2 presents the proposed content hierarchy. Section 3.3 looks at our proposed content development approach. Relevant e-learning pedagogies as well as recent e-learning developments (year 2006 - 2007) are included as concrete evidences to support and justify our approach. Section 3.4 presents our content development methodology. This chapter is summarized in section 3.5. Appendices A and B supplement this chapter.

3.1 Pedagogy Considerations for E-learning

Pedagogy is defined generally as the art or science of teaching and is anchored strongly at the heart of learning. It is focused on enabling learning and intellectual growth of students in contrast to instruction that treats students as the object of curriculum implementation. Successful learning pedagogy requires teachers to understand how students learn and must have the capacity and autonomy to design, implement and assess educational activities that meet the needs of individual students. E-learning pedagogy is one that incorporates this form of learning pedagogy but goes beyond it to include a deeper study into the incorporation of instructional strategies that take account of real-time personalized learning content-to-learner adaptability.

It is important to note that the proposed e-learning pedagogy is not intended to represent the full spectrum of complex tasks that comprise teaching; rather, it represents instructional activities that promote active student learning in the context of learning personalization. It focuses on the exploitation of information technologies for adaptation to the varying learning theories and diverse student needs.

Illustration of the ADDIE model (pedagogical aspects) and scaffolding framework is provided in Appendix A and Appendix B respectively. ADDIE is an acronym referring to the major processes that comprises the generic Instructional Systems Design (ISD) process: Analysis, Design, Development, Implementation, and Evaluation. It is one of the most popular instructional design models and has been considered by many e-learning professionals as the standard instructional model [Bruce, 2003]. This model is commonly used for developing training programs and has been documented, examined extensively and used successfully by many instructional teams over the years. The illustration is based on the Singapore Educational System and follows the Ministry of Education (MOE) learning guidelines (see <http://www.moe.gov.sg/>). Concrete learning content for teaching science courseware for secondary two students, namely Introducing Electricity and Household Electricity, is used as an example to showcase our pedagogical work.

3.1.1 Research Contribution – from Pedagogy Perspective

3.1.1.1 Gaps in E-learning Pedagogy Adoption

While e-learning has revitalized learning by transcending the boundaries of traditional learning, the use of computers per se for learning or teaching cannot make up for poor pedagogy and content. In relation to the new approaches to teaching, learning and the restructuring of organization training practices that e-learning brings, it is advocated that the current theories of pedagogy in this e-learning era requires an urgent review. Such a pressing need for the development of appropriate pedagogy for e-learning has been well documented [Ofsted, 1999; Lynch, 2001; Becta, 2002]. Although research effort has been carried out to develop appropriate pedagogy for integrating the use of technology into teaching, its development has lagged behind the massive investment in the provision of hardware, software and teacher training in using ICT [Newton, 2001]. Current processes of pedagogical changes associated with integrating educational technology, if any, seem to be evolutionary rather than revolutionary [Kerr, 1991; Cuban, 2001; Hennessy, 2005].

The pressing situation is not helped by the complex psychological nature of learning. A number of literature reviews on pedagogy also assert that models of learning pedagogy held by researchers and academics have become more complex over time (i.e. through the inclusion of cognition and meta-cognition [Watkins, 1999] or in terms of the complex exchanges between curriculum, assessment, pedagogy and the effect of these exchanges on student educational and social outcomes [Carr, 2000]). Also, the complex inter-dependent interactions between the teaching and learning process with the curriculum, assessment and pedagogical approaches further complicates the situation [Carr, 2000]. This together with the multi-faceted nature of knowledge and the mystical characteristics of its nature make an already demanding task even more difficult [Rowan, 2001].

Furthermore, most e-learning systems consist of several loosely integrated parts and functions (i.e. search engine, administration functions, ad hoc communications, etc.) that have no overall governing learning objectives. This couples with the total freedom and control for the learners' call for a sound learning pedagogy to guide not only the learners, but also the content providers (and in turn the learning materials). Thus, it is necessary that the e-learning system aids the content providers in constructing and structuring the learning material, using a learning theme that is self-explanatory, personalized and able to provide ample scaffolding support (in placed of 'direct' face-to-face advice and support).

Thus, aiming to build upon proven and tested pedagogical theories, we see a need to devise new pedagogical models that grow directly out of current e-learning practice or out of designs of e-learning curriculum materials. This view is also reflected by Kelly [Kelly, 2003] who sees the opportunity for a new pedagogy model to create stronger links between research studies and the practice environment, where presumably, learning materials are being used by students and teachers in realistically scalable ways. In particular, our focus is on new pedagogical ways of learning and organization that blossom from the inclusion of learning personalization in the e-learning environment.

Therefore, the aim of this research is to present our research findings, produce new scientific knowledge and generate innovative, yet theory-based pedagogical methodology in the field of educational technology. It is through this conservative form of research, based solidly on established theory and applying proven research methods, that we can add new knowledge incrementally and cumulatively [Wilson, 2005].

3.1.1.2 Contributions from Pedagogical Research

The pedagogical research contributions are:

1. The provision of a clear link between the theoretical principles of behaviorism, cognitivism and constructivism theorems, the practice of instructional design and the practice of teaching and learning in an e-learning context
2. The systematic translation of e-learning activities based on the psychological principles of learning
3. The characterization of essential learning activities using our **4-Ts** model (see section 3.3.2.1)
4. The characterization of scaffolding requirements in an e-learning context

3.1.2 Learning Theories Adopted

The first key outcome of our research is to offer a mapping of proven theoretical practices and theories onto our proposed e-learning model. While we emphasize on the three broad learning theories (chapter 2, section 2.1.8), other theories dominating our research approach come from Reigeluth's theory about content selection and sequencing [Reigeluth, 1983], Merrill's content presentation theory [Merrill, 1983], Gagne's

instruction lifecycle [Gagne, 1992] and Sweller’s theory about the cognitive impact of instruction [Sweller, 1988]. These theories together with specific portion of other learning theorems will be discussed as and when they are employed in the methodology.

We also devise learning conditions that are specifically designed to improve learning. These conditions are based on the behaviorism, the cognitivism, and the constructivism view. These learning theories, while not mutually exclusive, imply a different set of priorities for the learning and teaching practice. Here, we present the key learning priorities that were adopted from these three theories. However, before we proceed, it is important to clarify that while it is important to ground the research in a strong theoretical base, the learning theories do not answer any specific questions of learning nor do they provide any clear-cut solution to adopt. Instead, these learning theories offer focus, clarity and direction throughout the instructional design.

“Theoretical tool, in and of itself, is not an instructional design theory but defines instructional components that can be used to define instructional prescriptions more precisely”

Merrill [Merrill, 2001, p. 294]

“(learning) theories do not give us solutions, but they do direct our attention to those variables that are crucial in finding solutions”

Merriam [Merriam, 1999, p. 250]

Thus, the understanding of the theoretical framework and its incorporation within the scope of e-learning is important to ensure that the learning materials are effectively prepared and presented for learning personalization.

3.1.2.1 Implications of Behaviorism on E-learning

3.1.2.1.1 Theorems to be Adopted

Behaviorism theories impact e-learning design through the development of learning objectives. Learning objectives are generally statements depicting expectations on what the students are expected to know and understand after they have completed a learning experience. Learning outcomes, such as a culminating activity, product, or performance that can be measured, are usually associated with the learning objective. From the behavioral perspective, learning objectives take the form of precise statements that answer the question, *“What behavior can the learner demonstrate to indicate that he has mastered the knowledge or skills specified in the instruction?”* [Morrison, 2001, p. 91].

Such statements, while challenging to write, offer the content designers clear, measurable goals to which the content materials can be structured and designed. Mager [Mager, 1984, p. 21] further refines learning objectives to include the elements of (1) performance, (2) conditions, and (3) criterion. In summary, behaviorism's focus on objectively observable behavior is translated directly onto a concept of learning that must be based on the mastering of a set of predictable and therefore reliable behaviors. These behaviors, they advocate, will lead to desirable and demonstrable skills.

3.1.2.1.2 Examples in an E-learning Context

Typical examples of behaviorism-based instructions that can be adopted for e-learning are scenarios or problem-based learning. However, these materials must be separated into several smaller chunks of learning materials. The decomposition of the learning materials is essential to facilitate the ease of behavior reinforcement. Besides limiting the amount of information to be learned at each educational setting, the smaller chunk of learning material also allows frequent assessments and repetitions. However, due to the nature of such instructions, strict navigation over the learning contents must be enforced. Multimedia, interactivity and some of motivational factors are usually included as the need for frequent navigations and behavioral reinforcements might be detrimental to the learner's motivation and focus.

3.1.2.1.3 Pitfalls to Avoid

Behaviorism sees learning to be a finite and a reactionary process to an environmental condition. Therefore, it is more difficult to adapt the learning materials to accommodate the changing needs and demands of the knowledge economy. Hence, behavior-based materials heavily depend on content developers to create new environmental stimuli/condition whenever a 'new' behavior is required. Often, this is costly and time-consuming. Furthermore, behaviorism requires the behavior to be constantly reinforced, as behaviors that are not reinforced are likely to become less frequent and may even disappear [Merriam, 1999, p. 252].

3.1.2.2 Implications of Cognitivism on E-learning

3.1.2.2.1 Theorems to be Adopted

Cognitivism theories impact e-learning design through the development of task and learner analysis. Cognitivism places great emphasis on conceptual development through the integration of new information into the learner's existing knowledge in memory. Instructions must be based on the learner's existing mental structure, or schema to be effective. Hence, it is imperative to thoroughly analyze and consider the appropriate tasks needed for the learners to effectively and efficiently process the information received. Learner's characteristics must also be taken into account. As Blanton [Blanton, 1998, p. 173] states: *“the (instructional) goals should include learner needs and interest, reflect the concerns of society, and make every effort to insure that goals are focused at least toward the present and, hopefully, toward the future needs of the learner”*.

More importantly, although cognitivism does not inhibit the design of instruction like behaviorism, it does shift the focus of learning from the content developer-centric perspective to that of the learner-centric perspective. Hence, cognitivism-based instructions must be designed to emphasize the active involvement of the learner in the learning process such as learner control or metacognitive training [Flavell 1976].

3.1.2.2.2 Examples in an E-learning Context

Typical examples that can be adopted for e-learning are self-directed learning approaches such as taking control over the learning process (including progress monitoring), the usage of mind-maps (or concept maps), and learning content that emphasizes the recall of prerequisite skills. Such instructions are often coupled with practice items that are arranged with feedback so that the new information can be effectively and efficiently assimilated / accommodated within the learner's cognitive structure.

3.1.2.2.3 Pitfalls to Avoid

However, cognitivism places too much reliance on prior knowledge. While prior knowledge aids meaningful learning, the learner, brought up in a cognitivism setting, will be at a disadvantage whenever relevant prerequisite knowledge does not exist. Furthermore, the accurate elicitation prior knowledge still remains a key challenge.

3.1.2.3 Implications of Constructivism on E-learning

3.1.2.3.1 Theorems to be Adopted

Constructivism theories impact e-learning design through the development of task, learner and context analysis. The first two activities of task and learner analysis are similar to that of cognitivism as constructivism also places great design considerations on the learner's prior knowledge, understanding and interests. However, as opposed to the objective approach that cognitivism and behaviorism asserts, constructivism takes a more open-ended approach as each learner relies on his own mental models to accommodate new learning experiences. As such, learning is not easily measured and may not be consistent for all learners. Therefore, context plays an important aspect from the constructivism point of view [Spector, 2000]. Besides the context of learning, the context of the learner such as goals, perceived utility of instructions and perception of accountability [Morrison, 2001] must also be addressed.

3.1.2.3.2 Examples in an E-learning Context

Typical examples of constructivism-based instructions that can be adopted for e-learning are often collaborative in nature. It seeks to exploit the different impacts that the same learning topic can have on each individual student as a learning aid. Collaboration features such as team building with shared workspaces and group annotations are often included as part of the learning aid. Open-ended instructions (i.e. some basic introductory materials with simple guidelines and hints on how and with what means to start) are used to start the knowledge acquisition process. Self-analysis and collection of evidence to support the ideas or reformulate ideas in the light of new knowledge is encouraged.

3.1.2.3.3 Pitfalls to Avoid

However, the emphasis of individual perception of learning materials may lead to an instructional problem as there could be difficulties in evaluating the learning outcome.

3.2 Learning Resource Structure

The size and scope of the learning resources is a key pedagogical consideration. Besides the issue of network delivery [a larger file size will require a larger network bandwidth], the learning resource structure also affects the pedagogical nature of learning. For example, if the learning course contains only a few large-grained learning resources, then re-sequencing to form a new course to support a different pedagogical/personalization

approach may not be possible. On the other hand, although smaller grained learning resources are more flexible, the assembling of many fine-grained learning resources may require a considerable effort to tailor the narrative flow of the course to make sense in the context of the learner's learning.

While many researchers are proposing defining granularity in terms of the learning resource's file size or semantic density (as defined by LTSC Learning Objects Metadata Working Group [LTSC, 2000]), estimated learning time (Wisconsin Online Resource Center, [Thompson, 2005]), or complexity level per learning setting, such parameters are inappropriate. This is because no two learners learn at the same pace or perceive things in a similar fashion. Hence, the question is, what criteria can we use to generalize the learning time or complexity level for a group of learners? Furthermore, the specific learner profile is often not available at the conception of the course. Hence, such a granularity assessment using a learner profile is difficult to justify.

Here, we propose designing learning resource granularity based on learning concepts. A learning concept is the most basic form of learning and is used to convey paradigms of information. It covers concepts (What is a Router), facts (Water is made up of hydrogen and oxygen atoms) and principles/laws (Newton's Three Law of Motion). It is a fundamental learning concept that must be mastered first before more complex learning events or outcomes [i.e. procedure (How to set up a router), process (How traffic flows in a network)] can be learned.

A five-level content hierarchy is proposed.

1. Reusable Information Object (RIO)

The smallest grain of learning resource stored in the repository is termed a RIO. This is a chunk of information built around a single learning concept. The structure of a single RIO will be made up of the following compulsory items: a preview item, practice items, assessment items and content items.

2. Reusable Learning Object (RLO)

A RLO is a meaningful collection of RIOs that is built around a single learning objective. The structure of a single RLO will be made up of the following compulsory

items: an overview item, a summary item, assessment items, content items (5 ± 2 RLOs) and a governing learning objective item.

3. Topic

A topic teaches a high-level performance objective that is supported by a collection of related objectives. A topic is usually the smallest grain of learning recourse to be delivered. A topic can be mastered within a single learning setting (the duration of the learning setting is generalized using the expertise and detail level). The structure of a single topic item will be made up of the following compulsory items: an overview item, content items (5 ± 2 RLOs) and a governing performance objective item. Assessment items that test combinations of learning objectives are optional.

4. Module

A module has a single broad statement that clearly states the intended learning outcome(s) that must be achieved once the learning is completed. It combines several topics and usually cannot be mastered within a learning setting. The structure of a single module item will be made up of the following compulsory items: an overview item, content items (5 ± 2 Topics) and a governing learning outcome statement. Assessment items that test combinations of performance objective are optional.

5. Course

A course is the largest-grained of learning resource. It combines several modules and cannot be mastered within a single learning setting. The course structure is flexible.

3.3 Proposed Content Development Approach

This research marries the power of computer-based technology with an understanding of the psychological principles of learning to improve the educational outcomes. Incorporating our research work in using concept maps as an envelope for learning resources, we devise a novel and comprehensive content development methodology.

3.3.1 Overview of Approach

The proposed content development methodology will follow the phasing approach that is proposed by ADDIE Model - Assess, Design, Develop, Implement, Evaluate. In order to employ ADDIE Model's phasing approach, vital e-learning requirements such as learner-centric, concept-map based and personalized aspect of learning must first be modeled

into design tasks. After the design tasks have been conceptualized, they will then be sequenced, storyboarded and incorporated into the five phases of ADDIE Model. Importantly, the design tasks are modeled based on following key pedagogical strategies.

1. Instructional strategies (micro-level)
2. Learning events classification (macro-level)
3. Scaffolding framework

Additional key learning personalization issues that are considered includes:

1. Learners' learning preferences
2. Nature of knowledge delivered
3. Learning outcomes and learning/performance objectives
4. Prerequisites knowledge

The following sub-section details the approach and provides concrete evidence to justify their inclusion into our content development model.

3.3.1.1 Approach and Justification – ADDIE Model

ADDIE Model is a proven and well-known phasing approach for instructional materials [Plomp et al., 1992]. It is seen by learning pedagogists as the general instructional design model [Baruque et al., 2003], the best known design model [Gustafson, 1991; Carliner, 2000; Siemens, 2002] and is frequently used in academic circles [Siemens, 2002]. Its phasing approach is adopted by many learning pedagogists and organizations such as the Advanced Distributed Learning (ADL), the Idaho State University (ISU) College of Education, the Centre for Instructional Technology (CIT) at National University of Singapore (NUS), the Centre for Research in Pedagogy and Practice at National Institute of Education (Singapore), The University of Texas at San Antonio, etc.

In the e-learning context, ADDIE Model is considered by many e-learning professionals as the standard instructional model [Bruce, 2003]. Some examples of current practical learning solutions that have been developed in line with the ADDIE Instructional Design Methodology include learning objects application [Harman, et. al., 2007], expert system application [Tzeng, et. al., 2007], instructional design modeling [Najberg, 2007], and on-line e-learning instruction for designing learner activities, styles, and delivery systems [Idaho State University (ISU) College of Education, Science, Math and Technology

Education]. Its phasing approach has also been applied in e-learning 2.0 [Karl, 2006] and its learning advantages have been retested and validated through numerous recent test cases and learning settings [Nadolski, et al., 2007; Drozdova, et al., 2007].

3.3.1.2 Approach and Justification – Instructional Strategies (Micro-Level)

Instructional strategies determine the approach a teacher may take to achieve learning objectives. Instructional methods are often used by teachers to create the learning environment and to specify the nature of the activity in which the teacher and learner will be involved during the lesson. Instructional materials that have been effectively designed with sound instructional strategy will facilitate the desired learning outcomes for the students, enabling them to acquire higher order skills to be able to think and apply what they have learned in a different context. E-learning materials belong to this category of instructional materials. However, e-learning programs can, but do not always, deliver improved learning outcomes. This is due mainly to the fact that the current instructional strategy for e-learning, if any, is often adopted from traditional instructional strategy where face-to-face communication is assumed. It is advocated that in the absence or reduction in face-to-face communication, effective design of the e-learning materials has to rely on improved instructional design processes to reflect and replicate the real time interaction. Also, the e-learning materials need to be structured and personalized in a way such that it can communicate the knowledge in ways that enable students with diverse learning styles to understand and apply the knowledge that has been learned.

Our proposed instructional strategy is centered on how the learners are going to learn and chart their personalized routes to achieve their learning objectives. However, it goes beyond the basics of content sequencing or simple course structure generation. It is concerned with how the learners will interact and learn from the instructional content. When traditional instructional strategy is used in the context of e-learning, additional conditions must be considered to take into account those considerations that are traditionally deferred when face-to-face delivery / technologies limitation is assumed.

3.3.1.3 Approach and Justification – Learning Events Classification (Macro-Level)

Learning event classification must be carried out once the essential activities have been performed. Such classification is important as different learning events must be designed with different teaching and scaffolding methods.

The micro-level of instructional strategy prescribes four essential activities that must be performed. However, at the macro-level, different strategies should be employed according to the different context-dependent learning events involved. For example, when the courses to be taught requires only straight-forward knowledge impartation (i.e. rote learning where memorization is a key necessity for learning), the teaching strategy is usually centered on the use of the storage-retrieval of the knowledge process. As for courses that teach problem solving, this usually does not involve the gathering of new knowledge but may involve teaching strategies that help the student in the reorganization of their cognitive data or remembering how to deduce or apply knowledge so as to achieve required solutions.

Hence, different sets of guidelines should be provided for the different classifications of the learning events. These learning events are classified based on their ability to promote the transfer of knowledge. While the learning events are set in various settings, it is important that all these learning events must start from the learner's perspective [Simons, 1999]. The learning events also take into account the target audience (conducted under the Analysis Phrase: Learner Profile Analysis) by adopting the research findings from Simons' Transfer of Learning [Simons, 1999]. i.e. for young students, design the learning materials such that the students apply their prior knowledge more actively, overcome some of their pre-conceptions and help them to learn on their own (see scaffolding section); for adults, design the learning materials through embedding a learning to learn approach into their regular training so that they can learn how to transfer their knowledge.

3.3.1.4 Approach and Justification – Scaffolding Framework

Scaffolding is seen as one effective instructional means to escalate one's understanding from a novice viewpoint to that of an expert. Scaffolding instruction as a teaching strategy originates from Lev Vygotsky's sociocultural theory and his concept of the Zone of Proximal Development (ZPD). The zone of proximal development is the Vygotskian

concept that defines development as the space between the child's level of independent performance and the child's level of maximally assisted performance [Vygotsky, 1978; Bodrova, 1996]. This concept of ZPD was later broadened by contemporary Vygotskian scholars to serve as a general metaphor for human development in a sociocultural context (e.g. [Newman, 1993]).

The term “scaffolding” was coined by Burner [Wood, 1976] to specify the types of assistances that make it possible for learners to function at higher level of their zones of proximal development. The term is often used to describe how an expert can facilitate the learner's transition from assisted to independent performance [Meyer, 1993; Berk, 1995]. With adequate support, this novice to expert transformation [Quitana, 2004] can enable the student to create meaningful inferences from seemingly unrelated raw data. Following the guidance of the more knowledgeable expert, the student will become competent with academic tasks that are initially beyond their ability [Wood, 1976; Palincsar, 1998]. The more knowledgeable expert can also successfully diagnose the complex needs of students at various stages of the intended learning and employ proper instructional strategies adaptively to their progress [Tabak, 2004].

Scaffolding differs from other types of instructional strategies through the key characteristics of fading and student support. Fading refers to the gradual reduction of support by the more knowledgeable agent in successful tutor-tutee [Wood, 1976], mother-child [Wertsch, 1985], teacher-student [Fleer, 1992; Flick, 2000] or expert-apprentice relationships [Brown, 1989]. Scaffolding-minded learning meddling can incorporate fading either as an implicit part of student's learning or as an explicit part of an active intervention strategy. While fading is an important aspect of scaffolding, the detailed mechanisms of fading in the scaffolding framework are not clearly understood [Stone, 1998].

Besides the basic teaching advantages that scaffolding brings, the unique characteristics of e-learning requires scaffolding teaching strategy to be selected as one essential component of e-learning pedagogy because of the following reasons:

1. Scaffolding can provide individualized support based on the learner's ZPD [Chang, 2002]. Individualized or personalized support is of utmost importance to e-learning and has been discussed in details in the previous chapter.
2. The scaffolds enable the facilitation of a student's ability to build upon prior knowledge and to internalize new information [Van Der Stuyf, 2002]. Prior knowledge activation has been highlighted in the earlier section (4-Ts of a learning session) as one of essential e-learning methods.
3. The use of fading can mimic the process of teacher-student interactivity. This is important because in a traditional learning environment, the teacher is in constant communication with the students and thus, is sensitive to the progress of individual students. Hence, the teacher, equipped with ample knowledge of the student, can select the most effective fading mechanism for guiding the different students in completing complex reasoning tasks. However, when learning is housed in an e-learning context, the luxury of face-to-face communication and real-time teacher-student assessment is no longer available. In the absence of such teacher-student interactivity, it is important to address such complex reasoning tasks early and incorporate it into the instructional design.
4. The scaffolds that usually consist of models, cues, prompts, hints, partial solutions, think-aloud modeling, and direct instruction [Hartman, 2002]) can be easily programmed and delivered over the network.
5. Noting Vygotsky's view that the learner does not learn in isolation (in fact, learning is strongly influenced by social interactions which take place in meaningful contexts) and the ironic fact that the 'distance aspect' of e-learning while transcending the boundaries of traditional learning, has minimized (or eliminated) social interaction, scaffolded instructions, by its nature, can be designed to overcome this [Geow, 2001].
6. Scaffolding's, inherent disadvantages (i.e. time consuming to develop supports and scaffolded lessons to meet the needs of individual, most teachers are not properly trained to implement scaffolding instructions, most teachers are not trained or comfortable to give up some of their control and allow the students to make errors and traditional teacher's education does not teach scaffolding as a teaching strategy) as pointed out in [Van Der Stuyf, 2002] can be solved through the use of e-learning

technology. For example, as e-learning materials emphasized the characteristic of reusability, the scaffolds can be developed once and reused or repurposed many times. Hence, the scaffolding development process can no longer be considered time – consuming as the return in investment (ROI) is manifold. Also, the teacher's inexperience and reluctance to build scaffolds can be minimized or even solved through the semi-automation of the development process and the incorporation of a scaffolding guide.

In view of the advantages that scaffolding can bring to e-learning, this research proposes an instructional design which includes a fading approach that enables the curricular materials to fade scaffolds systematically based on the diagnosis of student progress

3.3.2 Design Tasks

With the key pedagogical strategies identified, the content development methodology then prescribes the design tasks that model the individual learning pedagogical activities.

3.3.2.1 Instructional Strategies – 4-Ts

The instruction strategy employed at the micro-level requires key activities to be performed regardless of the kind of knowledge to be imparted. These key activities are essential to the act of learning and are termed as the **4-Ts** of a learning session: Target, Training, Transfer, and Transformation. The 4-Ts are centered on the three main learning concept of (1) identification of goals, (2) identification of critical factors required to achieve the goals and (3) determination of how achievement can be measured. In order to achieve these three main concepts, a total of four activities must be performed.

3.3.2.1.1 4-Ts of a learning session: Target

Main Goals:

- Identify the *Governing Motivation* underlying the need for training
 - Educational-centric: educate/inform
 - Performance-centric: Specific skills acquisition
- Identify the *Entrance Pre-requisite List*
- Identify the *Non-Coverage (Exclude) List*

The target (goal) identification is conducted at the conception of a new course. It is situated at the heart of the pre-instructional activities. The learning target identification is important because human beings have a built-in goal seeking "success mechanism" that is part of the subconscious mind [Maltz, 2002]. This success mechanism is constantly finding ways to help us reach our targets and find answers to our problems. According to Maltz [Maltz, 2002], learners work and feel better when their success mechanism is fully engaged going after clear targets. Moreover, setting the learning target also helps to concentrate the effort, establish priorities, and to provide a development roadmap for the learning materials. Thus, this phase explicitly marks the generating of the governing motivation underlying the need for training. It also states the essential questions that must be thoroughly examined before the commencement of the course design and includes defining the essential entrance criteria as well as the non-coverage list.

In the SME perspective, the most important step in any learning event is to focus on guiding the learners to understand the learning concepts. Depending on the desired type of learning outcomes, different teaching styles and hence, their appropriate learning (teaching) events must be adopted. Hence, the first step in any instructional strategy is to state clearly the underlying motivation behind the learning event. The motivation can be categorized under two teaching perspectives: educational-centric or performance-centric.

1. Educational-centric Motivation

The main purpose in an educational-centric learning event is to educate or inform. Its sole purpose is to augment knowledge and is inclined towards imparting theoretical and abstract concepts. This mode of learning is extremely beneficial in conveying paradigms of thinking and information. The application of knowledge and the development of communication skills are however secondary. Such learning events are characterized by learning contents that are hierarchically organized and aim to guide the learning process through structured syllabi and tests. The training is usually housed in a certain context but the students are assessed both inside and outside the arena in which they hope to minister. Upon the completion of the training, the newly acquired expertise which attests to the level of training will be recorded in learner's cognitive map (the learner's

cognitive map is our novel method of eliciting the learner prior knowledge (see chapter 4, section 4.3.1)).

2. Performance-centric Motivation

A performance-centric learning event on the other hand looks at the practical usage of the knowledge. It focuses more on the application of the knowledge rather than its fundamentals. This mode of training is based on the premise that students learn most effectively through experiences and practice in a deliberately organized program. It uses real life examples as the basis for purposeful training. Such a mode of learning is highly relational yet unstructured in the sense that training is focused on working towards problem solving and the teaching materials are structured outside the normal school curriculum. The student has complete control over his learning and this mode is participatory in nature.

The formulation of the learning/performance objective together with the non-coverage list is a key aspect of the instructional strategy that caters to the learner-centric aspect of e-learning. The learning/performance goal and the non-coverage list states, in less formal terms (so that the learners can understand), what is or is not required of the learners during the learning event. This is essential as the learners need to be informed of what is going to happen during the learning event so that they can focus and determine what is relevant to them and what is not.

The entrance pre-requisite list sets the context and essential pre-requisites. This list connects the learners with the training tasks and helps the learners to assimilate the new information into a context, based on what they should already know. Hence, this list sets the foundation on which the present learning concept will build on. This is a necessary step to aid in the synthesis of knowledge into the learner's cognitive structure.

3.3.2.1.2 4-Ts of a learning session: Training

Main Goals:

- Activation of prior knowledge
- Identify the nature of the Training Tasks
 - Collaborative / Individualistic

- Analytical / Passive
- Divergence / Convergence
- Discussion / Questioning
- Project / Assignment
- Demonstration / Practice
- Lecture / Case Study / Role Playing
- Identify the Training Materials
 - Main subject matter – purely on textual content
 - Supplementary subject matter – purely on textual content
 - Subject matter presentation
 - Learning styles
 - Cognitive level
 - Interactive multimedia
- Identify the Training Support
 - Subject Matter Sequencing
 - Chronological
 - Cognitive level
 - Knowledge Gap (only if profile of learner is available)
 - Taxonomic (based on structure)
 - Problem / Case study-centric
 - Scaffolding Means
 - Degree of Abstraction
 - Degree of Complexity
 - Degree of student independence
 - Pace of learning

With the target set, the pedagogical consideration now falls on the instruction set. However, before any content deliberation, in a learner-centric learning environment, it is crucial to invite the learners to clarify where they stand, at present, in terms of the new content. The prior knowledge activation task is important. Many prior studies have demonstrated the importance of informal and formal prior knowledge (see [Dochy, 1992]). Essentially, prior knowledge activation sets the stage for learning by sharpening

the perception of the learner. It tells the learner not only what he has to learn but also what the person perceives he already knows. Sometimes this is a rude awakening, sometimes a corroborating experience. The result of such reflection when compared with the system record of the learner's past expertise can aid in the identification of the appropriate starting point of his learning route. The activation of prior knowledge is being applied as a sort of learning preparation. It is used to connect the new knowledge with the existing ones; to synthesize prior knowledge with the new content. Hence, the main pedagogical consideration here is to effectively plan the querying methods to accurately retrieve the learner's prior knowledge and investigate how to connect the learner's prior knowledge to the new content. Once the prior knowledge correlation is achieved, then training can begin. The training development plan is sub-divided into 3 phrases: Training Tasks, Training Materials and Training Support.

3.3.2.1.2.1 Training Tasks

Training Tasks identify the nature of the content presentation. The training task defines the way to structure the training. Depending on the type of learning event, different training tasks are used to create and ensure engagement of the learners with the new content. The structuring of the nature of the training task is one way to effective teaching as it dictates and creates a consistent overall design for the training. The teaching challenge here is not to present the training materials as static content but as an integral part of learning where the learners have to work over, contest, digest and recreate the new information to fit into their cognitive structure. Instead of receiving information, the aim here is to turn the learning process around and offer an opportunity for the students to construct their own knowledge through the use of concept mapping techniques.

3.3.2.1.2.2 Training Materials

Training Materials identify the type of content. It is important to note that no form of content presentation, formatting or delivery medium is taken into consideration. While SME has to deliberate over the type of learning styles, cognitive level and interactive multimedia for the learning content, it is the pedagogic issue that is of concern, not the technological or delivery aspects. The full content structure need not be formulated here. Only the key concepts and its associated teaching method are devised.

3.3.2.1.2.3 Training Support

Training Support is to ensure that the learners get the most out of training materials. Depending on the nature of the teaching tasks and the type of learners, different supports such as sequencing, pacing, scaffolding, complexity, and abstraction level will be employed. Training support, particularly scaffolding, will be discussed in section 3.3.2.3.

3.3.2.1.3 4-Ts of a learning session: Transfer

Main Goals:

- Pre-assessment
- Practice
- Post-assessment
- Remediation
- Enhancement – introduction to next level of understanding

The transfer of learning is the most important aspect of the learning. Transfer of learning is the application of skills and knowledge learned in one context being applied in another context [Cormier, 1987]. The transfer of learning is important as the learning context is often different from the context of application. Hence, the goal of learning is never accomplished unless transfer occurs. Successful transfer of learning requires (1) training content be relevant to the task, (2) the learner must learn the training content, and (3) the learner must be motivated.

The triggering of knowledge synthesis usually lies in the activation and application of the newly acquired knowledge. Such a knowledge transfer can occur in three ways: (1) from prior knowledge and skills to new learning [e.g. a student from China who reads a document in English but uses the knowledge and skills from his native language (Chinese) to interpret the contents], (2) from new knowledge and skills to new learning situations – learning now preparing for later learning [e.g. a student who is learning the concept of differential equations but knows that he has to master the concept of stability later (the concept of differential equation is an essential pre-requisite to the learning of the concept of stability)], and (3) from new knowledge and skills to applications in work and daily life – learning for practice (e.g. a student from China who learns to speak English in the

lesson, knowing that he has to perform it in real-life context when he speaks to his professors) [Simons, 1990].

The transfer of knowledge can also be classified under near or far transfer. In near transfer there is a close connection between the learning situation (or the prior knowledge) and the application (or the new learning situation). In far transfer the distance between prior knowledge or learning and application (or the second learning situation) is much greater (see [Mayer, 1972]). These modes of knowledge transfer will be catered for at the macro-level of the instructional strategies.

In general, in order to demonstrate that learning has taken place, practice and assessment are included in the transfer of learning. However, it is important not to confuse practice with assessment. Learners at the practice stage are still learning; they are not being assessed. There are many different learning strategies for demonstrating understanding of the course ranging from requiring the learners to compare, classify, induce, deduce, analyze, construct or to make abstraction of the new acquired knowledge. Typical activities include open-ended questions, summaries, quizzes, and assignments.

While both practice and assessment are classified under the transfer of learning, the assessment of learning forms the evaluation portion of the learning event and must be based on a well devised methodology to determine if the e-learning event has been successful and learning has taken place. If the learners did not achieve the required understanding of the subject matter, sufficient scaffolding should be incorporated. Beyond the application consortium of practice and assessment is the higher level of understanding: remediation and enhancement.

Remediation is the application of the newly acquired knowledge in an entirely new situation. This gives the learners an opportunity to transfer what they have learned to other situations and use it in different ways. Collaborative learning is one good method of ensuring that the learners apply what they have learnt to new situations. Lastly, the enhancement aspect looks to bring the learners to another cognitive level. It introduces the ‘post-requisites’ concepts and broadens the understanding.

3.3.2.1.4 4-Ts of a learning session: Transformation

Main Goals:

- Assess the accountability of the teaching
- Feedbacks – both learners and teachers
- Revision

Transformation is the final phrase that assesses the change in one's approach that is brought through by the training and transfer. It sets the accountability of the training process and is also the channel through which the learning content is enhanced and revised by appropriate reviews and feedbacks. Feedback and comments may be in the form of electronic surveys or direct feedback from the teachers or learners.

Transformation differs from transfer in the sense that transfer of knowledge is more of an application of the knowledge but transformation is set at a higher level as it looks at the behavioral change that is brought about by the new training.

3.3.2.2 Learning Events Classification – Educational and Performance

The learning events can be classified under two main categories: educational learning events (What-is learning events) and performance learning events (How-to learning events).

3.3.2.2.1 Educational Learning Events

The educational learning events are typically theoretical in nature and cover all fundamental learning theories and concepts such as facts, laws and principles. It must be delivered at the conception of the lesson and form a bulk portion of the learner's initial academic learning phase. It is characterized typically by knowledge impartation (from teacher) – cognitive awareness (student) – cognitive reorganization (student).

3.3.2.2.1.1 Factual Learning

Factual learning events are characterized by learning materials that are presented as a hierarchy of topics. One topic is set as the main topic and all the other topics are termed as the sub-topics. These sub-topics are either pre- or post-requisites concepts of the main topic and must bring the learner towards comprehending the main topic. Learning is structured, logically and usually sequenced in a linear fashion where there is an explicit

assignment of topics as the entrance and termination learning points. Learning always proceeds from the known to the unknown and is directed by the teacher. It involves directing the student attention towards specific learning (main topic) in a highly structured learning environment (main topic – sub topics interlinkage). The topics and contents are usually broken down into small modules, taught and assessed individually through explanation, demonstration and practice. The attention of students is important and listening and observation is the key to success in this phase.

3.3.2.2.1.2 Conceptual Learning

Conceptual learning events are characterized by learning materials that can be presented as verbal, written summary, outline of a topic or visually as a set of concept maps. It can exist as a new concept or tied to a Factual learning concept (either at the beginning of the unit, or module).

While conceptual learning is also centered on a main topic, the main learning objective is not to master the main topic but to present the big picture that houses the main topic. Gaining an overview of how the new concept is related to other concepts is another form of learning by itself. Besides mastering the fundamental theories underlying a particular concept, the ability to link and synthesize new knowledge into one's existing cognitive structure is another complex learning process that requires supervision. In addition, connecting new ideas to information students already understand makes it easier to retain. The topics that are presented in the map need not be pre- or post-requisite concepts of the main topic. Instead, these topics can be stand-alone main topics by themselves that aim to portray the association between themselves and the main topic.

Learning in conceptual learning events is usually unstructured in the sense that the learners are allowed to explore all the relevant information in a manner that is comfortable for them. This form of learning is similar to learning from the Web where learners can follow any relevant keywords or concepts to look up a topic. Conceptual learning allows the learners to explore any concept quickly and easily. It also enables the learners to see how one concept is related to other concepts. Conceptual learning typically distills difficult or complex concepts through the use of simpler definitions or

examples before showing how such a concept relates to other information. It emphasizes the learning process of ‘cognitive reorganization’ to make a concept meaningful.

3.3.2.2.1.3 Supervised Learning

Supervised learning events are characterized by a substantial amount of teacher/agent help. In this kind of learning event, the learners undertake various tasks with the help of a mentor. Help can take the form of direction intervention to scaffolding to learning hints.

While complex tasks form the bulk of supervised learning events, supervised learning events need not be characterized fully by the nature of complexity. Tasks/theories that are ambiguous, subjective or open-ended in nature can also benefit from having a mentor’s supervision. The structure of supervised learning events will vary depending on the type and intensity of supervision being employed.

It is however important to note that this type of learning event should only be used when needed, so as not to make the learners over reliant on the mentors’ help.

3.3.2.2.2 Performance Learning Events

The performance learning events, in contrast to educational learning events, take a more practical approach to learning and typically cover handy skills that are required to perform some specific tasks. It moves from the conception of the learner as a student to one of the learner as a professional who needs to gain and apply the knowledge so as to contribute to the field. Hence, such learning modes usually rely on authentic and diverse real life problem-based examples. It focuses on helping the students to link and interconnect their learning. Opportunities are provided by such events to break away from the compartmentalization of knowledge and skills and to help the students to construct a better network of knowledge and skills. Furthermore, such learning events are an authentic form of learning that prepares the students for the increasing demands of the workplace [Kwok, 2004].

Performance learning events, when used in a corporate context, is like on-demand training (the latest trend in e-learning). Companies using e-learning technologies have found that long lectures do not cut it online. Short, targeted learning segments with

simulation or how-to scenarios let employees take classes when they have time or when they need the help. It is far less disruptive than taking week-long seminars. When the learning events are housed in an academic context, it is similar to project-based learning which is action-oriented and focuses on doing something rather than learning about something [Moursund, 1999]. However, project-based learning while important has many practical constraints (i.e. too much work to prepare, difficult to assess, learning outcomes undetermined, difficult to assess if learning takes place, no in-process feedbacks, cannot monitor and facilitate progress, etc.) that cause great difficulties in realizing the goals of project-based learning [Kwok, 2004]. Hence, we feel that the coverage of project-based learning is too broad and undefined. We advocate that through the proper classification of such project-based learning into performance learning events, we can divide and simplify the complex cognitive and meta-cognitive processes while maintaining the learning outcome. Also, through the usage of computing resources, many taxing (preparation works, grading of reports, progress tracking and monitoring of students) and ambiguous tasks (assessment criteria, learning outcome formulation) can be programmed.

3.3.2.2.2.1 Guidelines

Guideline learning events are characterized by learning events that are generally more process and procedure oriented. It typically teaches a flow of events that describe how something works. This set of guidelines is applicable in all scenarios and minimum or no interpretation on the part of the learner is required. However, although the tasks undertaken are the same, the learners have to be trained in understanding what and why such actions are undertaken and also in some troubleshooting techniques.

The most difficult aspect of such learning events is to generate interest in the learner as such learning events are typically less intellectually stimulating.

3.3.2.2.2.2 Reference

A reference learning event is one that is similar to the guideline learning event. However, as opposed to the guideline learning event where the process or procedure can be replicated in all situations, a reference learning event is one that requires sound judgment on the part of the learner. While a reference learning event contains the necessary

information on a particular problem domain, some discretion must be exercised when devising the solution to the problem.

3.3.2.2.3 Troubleshooting

A troubleshooting learning event is the most common type of informal learning where the learning event consists of a set of questions representing common problems. Instead of presenting the theory underlying the problems, the set of possible actions to the problem is presented in a form of a decision matrix. The possible scenario that will occur as a result of the action selection will be presented. From the set of actions taken, the learner is indirectly revealing his understanding of the problem. Thus, this type of learning event is an effective way of identifying the learner's misunderstanding about certain concepts.

3.3.2.3 Scaffolding Framework – Design Tasks

Goals	Tasks	Reasons
Curriculum Considerations	<ol style="list-style-type: none"> 1. Learning Curriculum Goals 2. Context 	The teacher must first identify the curriculum goals and select the context in which the learning concept resides. He must be knowledgeable of the content and be sensitive to the learners (e.g. aware of the learners' background knowledge and misconceptions) to determine if they are making progress.
Learner Considerations	<ol style="list-style-type: none"> 1. Needs Analysis 2. Knowledge Gap Analysis <ol style="list-style-type: none"> i. Identify the students' current knowledge point ii. Identify the students' target knowledge point iii. Identify the necessary teaching materials to move the learner from his current knowledge point to the target knowledge point 	<p>The teachers must identify the learners' needs in order to effectively plan the teaching routes and methods. This is an essential step to cater to the learner-centric aspects of e-learning. Learner consideration is an important element in the pre-engagement process that links the students with the curriculum. The pre-engagement process aims to establish a shared goal between the teachers and the students so that appropriate teaching plans can be selected to suit the students. It is through this shared goal between the teacher and the student that the student will become more motivated and invested in the learning process (and hence reduce the drop-out rates).</p> <p>The process of needs analysis aims to conceptualize the learner's goals and their immediate learning needs. Ideally, this process assumes that the profile of the learner is available at curriculum design time and the teacher is able to work with the student to plan the instructional goals and materials. However, in the context of e-learning, the distance aspect and the learning on the fly mentality makes the needs analysis difficult or even impossible. Even with the advancement of the networking technology, the exact profile of the learners who are going to take this course can never be established. Hence, this needs analysis is actually conducted based on the nature of the course offered and the feedbacks that are received based on similar courses.</p>

Chapter 3: Learning Personalization – A Pedagogical Content Development Methodology

		<p>For example, if a course on “Web Service” is to be designed, the needs analysis will be centered on the context whereby such courses will be held in (i.e. info-communication context audience). Based on similar courses held in that domain (i.e. software engineering), the learners’ profile and their feedbacks will be analyzed to identify and estimate the learning needs for that particular group of learners.</p> <p>Once learning needs are identified, the teacher must assess the knowledge gap. The gap analysis has three phases. Phase one identifies the learner’s current knowledge point. Without a profile of the ‘real’ learner, this phase relies on the teacher’s past experience when conducting such courses. He will have to assess what the learner is currently able to do without help. Basically, the pre-requisites are a good indication of the learners’ current ability and can be used as an effective starting point. Next, the teacher will identify the desired performance of the student (that is, the goal of the student which is identified in step 1). Lastly, the teacher will analyze what is to be achieved to move the learner from the current knowledge point to his desired knowledge point. This analysis will form the core of the next step.</p>
Scaffolding Design	<ol style="list-style-type: none"> 1. Scaffolding Path 2. Fading Design Tasks 3. Scaffolding level assessment 	<p>The scaffolding path reestablishes the starting and the finishing point of the scaffolding design. The starting point is the learner’s current knowledge point and the finishing point is the learner’s target knowledge point. The fading design will be classified into three stages: First, the teacher does it, second, the group or peer does it, and then lastly, the learner does it. (See scaffolding assessment level table for details (Table 3.7))</p> <p>Stage I: Teacher-directed learning</p> <p>This stage is the teacher-directed phase where the level of support used is the greatest. Most of the learning content will be covered in details and explicit links to concepts and references will be provided. Learning process is step-though and the teacher will model how to perform entirely new or difficult tasks.</p> <p>Stage II: Group / Peer-directed learning</p> <p>This stage is the group / peer-directed phase where level of support is reduced. Only about half of the major learning concepts will be covered in details as this phase relies on induced inference. This is semi-guided learning in the sense that help will not be as spontaneous as in stage I and teacher support will be asynchronous. However, there will be aids and implicit hints to new or difficult concepts and tasks to simulate the cooperation between the teacher and learner to perform the tasks together.</p> <p>Stage III: Learner-directed learning</p> <p>This stage is the learner-directed phase where level of support is negligible. That is, the learner is no longer dependent on the teacher’s extrinsic signals to begin or complete a task. Only key learning concepts will be covered in details as this stage emphasizes on intellectual prediction, concept synthesis and reflection. While this is</p>

		an independent practice stage where the individual learner can demonstrate their task mastery, this is still practice and not assessment. Hence, necessary aids will be provided if the learner needs it.
--	--	---

Table 3.1 Scaffolding framework – Design Tasks**3.3.2.3.1 Scaffolding Assessment**

	Level 1	Level 2	Level 3
Support	Full support	Moderate support	Minimal support
Direction	Teacher-centric	Group/Peer-centric	Self-centric
Coverage	> 70% Major Concepts	40~70% Major Concepts	< 40 % Major Concepts
Granularity	Small (Problem-based)	Moderate (Case-based)	Large (Scenario-based)
Assessment	Frequent (weekly)	Less Frequent (monthly)	Minimal (quarterly)
Discussion	Scheduled: 1. Mandatory group feedback 2. Mandatory peer feedback 3. Discussion boards 4. Forums 5. Emails	Unscheduled: 1. Mandatory peer feedback 2. Discussion boards 3. Forums 4. Emails	Unscheduled: 1. Discussion boards 2. Forums 3. Emails
Tutor's Aid	Synchronous/Asynchronous e.g. 1. Q&A sessions 2. Discussion boards 3. Forums 4. Emails	Asynchronous e.g. 1. Discussion boards 2. Forums 3. Emails	Asynchronous e.g. 1. Emails
Task Support	1. Provide explicit links to concepts 2. All references provided 3. Guided learning 4. Step through process	1. Induce inference building 2. Implicit hints to relevant concepts 3. Semi-guided learning	1. Intellectual Prediction 2. Concept reflection 3. Concept synthesis

Table 3.2 Scaffolding Assessment Levels**3.3.2.3.1.1 Justification of Scaffolding Assessment Levels**

1. Support /Coverage

This is in line with the scaffolding principles of fading.

2. Direction (Grouping)

Grouping of learners as the scaffolds reduce is of utmost importance as an essential element of scaffolding is that the participants must be in social interaction whereby they negotiate or compromise by constantly striving for a shared view of the situation

[Berk, 1995]. Following the use of teacher-provided scaffolds (level 1), level 2 introduces the usage of grouping to allow the students to ask and answer questions. In this type of learning environment, students help students in a small group setting but still have some teacher assistance. This is an essential step in the process of decreasing the scaffolds provided, and is needed by the students [Hartman, 2002]. Lastly, level 3 removes the group / peer support to let the student assume full control, and the problem-solving capabilities, knowledge and responsibility can be said to have been transferred from the teacher to the student [Berk, 1995].

3. Granularity

Supplying the learners with appropriate learning materials is an important tool. In starting small, the teachers relate how ‘calling it as they see it’ was very effective [Kearn, 2000].

3.3.2.3.2 Scaffolding Principles

The scaffolding principles originate from [McKenzie, 1999] but has been modified to suit the e-learning context. Also, some of the activities and tasks proposed by (Ellis et al., no date) has been incorporated into the scaffolding framework.

1. Begin with what the students can do

It is important to start the course with something that the learner can associate with. The activation of one’s prior knowledge is one excellent introduction to the lesson as it allows the learners to be aware of their strengths and to feel good about what they can achieve without help.

2. Clarify learners’ current knowledge point

This invites the learners to clarify where the learners are at present in term of the new concepts that they desired to master. This is used in combination with the “begin with what the students can do” principle to set the stage for scaffolding.

3. Begin with small, simple granular tasks

Small, simple granular tasks should be used at the conception of the course. Although the learner needs challenging work in order to learn, frustration and loss of focus will set in when he experiences a constant cycle of failure. Hence, it is important for the learner to experience constant success and also a sense of fulfillment and confidence

before he embarks on larger and more complex tasks. As the learners' mastery increases, the granularity and complexity of the tasks should increase progressively.

4. Frequent Assessment

It is important to know when the time to stop is. Scaffolding is important to help the learner perform certain tasks, but too much might impede learning and create over-reliance. In a traditional learning environment, the teachers need to watch for clues from the student that show when and how much teacher's assistance is needed. In the context of e-learning, frequent assessment of the learner's progress will be used in place of the teacher's judgment. Scaffolding support needs to be removed gradually as the learners demonstrate mastery and the support will be removed totally when the learner is able to perform the tasks independently.

5. Engagement of student with his learning

The learner will become more motivated and interested in the learning process when he is able to dictate and plan his instructional goals and learning route. Hence, to help the learners monitor their own progress, the teacher must assign regular checkpoints and guidelines such that the system can automatically summarize the learner's progress and explicitly note and record any particular behaviors that contributed to the learner's success or failure.

6. Use scaffolding only when appropriate

Not all tasks, be it complex or large, need scaffolding. Also, all learners learn differently. Hence, not all the learners need scaffolding. Appropriate surveying of learner's learning preferences and needs need to be conducted.

7. Practice generating more than one possible prompt

The first prompt or hint that the learner receives may fail. Therefore, more than one prompt or hint using different methods or cues is needed to generate the appropriate response. Typical tailored assistance includes cueing, prompting, questioning, modeling, explicit message box, mandatory discussion and request clarification. The teacher must devise teaching methods and plans to 'advise' the assessment agent to use them when needed or adjust them to meet the learner's needs.

3.4 Proposed Content Development Methodology

Legend:

Ť: Main Task to be performed by the content developer

T: Sub-task to be performed by the content developer

Ä: Task automatically performed by system without the need for human intervention

3.4.1 Analysis Phase

Steps	Purposes	Activities
1. Feasibility Analysis	Justify the creation of E-learning content	Ť Conduct Feasibility Survey Ť Analyze Feasibility Report
2. Pedagogy Analysis	Establish the type of instructional strategy	Ť Conduct the 4-Ts Ť Determine the type of learning event
3. Course Profile Analysis	Determine the scope, context, and performance augmentation	Ť Identify the overall (high level) educational Goal Ť Identify the conceptual scope Ť Identify the application domain Ť Identify the type of performance augmentation
4. Learner Profile Analysis	Establish the target audience and their learning needs	Ť Identify the target audience Ť Identify the target audiences' learning needs Ť Identify and categorize target audiences into their different learning styles Ť Perform the course-to-learner profile matching

Table 3.3 Analysis Phase**3.4.2 Design Phase**

Steps	Purposes	Activities
1. Performance Design	Craft the learning and performance goals into quantifiable objectives	Ť Identify all the measurable learning and performance objectives Ť Determine the assessment guidelines
2. Scaffolding Design	Defining the scaffolding approaches	Ť Identify the learning context in which the concept resides Ť Scaffolding Goal Design Formulate: T Scaffolding Aim T Scaffolding Problem Ť Scaffolding Tasks Design Formulate Tasks to address: T Activation of prior-knowledge T Encoding specificity – resemblance of problem to learning concepts T Elaboration of knowledge application Ť Learner Considerations Formulate Tasks to address: T Identify the learner's goal through the process of needs analysis T Knowledge Gap Analysis - Assess learner's current knowledge point (Assess what the learner is currently able to do without help)

Chapter 3: Learning Personalization – A Pedagogical Content Development Methodology

		<ul style="list-style-type: none"> - Assess learner's desired knowledge point (Find the desired performance) - Determine the level of learner support to design <p>(While all 3 levels should be designed, the most appropriate level should be determined and designed first and used as a guide to spearhead the design)</p> <ul style="list-style-type: none"> - Ā Knowledge Gap Mapping <p>(Analyze what has to be achieved to move the learner from the current knowledge point to his desired goal)</p> <ul style="list-style-type: none"> • Specific what is necessary to narrow or eliminate the knowledge gap • Specific what is the appropriate instructional strategy to scaffold the design • Determine how to measure success or failure
3. Course Structure Design (Extrinsic Properties)	Generate the course structure in terms of tree-hierarchy and concept map structure	<ul style="list-style-type: none"> • RLO Identification <p>Formulate:</p> <ul style="list-style-type: none"> • Educational Goal • Cognitive Level • Type of content presentation • Associated Keywords <ul style="list-style-type: none"> • RIO Identification <p>Formulate:</p> <ul style="list-style-type: none"> • Educational Objective • RIO type (Concept/Fact/ Principle/Law) • Associated Keywords <ul style="list-style-type: none"> • Course Tree Hierarchy • Course concept map • RLO/RIO search for reused/repurposed • Verification of course structure • Verification of RLO/RIO status (develop/reuse/repurpose)

Table 3.4 Design Phase

3.4.3 Development Phase

Steps	Purposes	Activities
1. Course Composition Overview Development (Intrinsic Properties)	Generate the course composition in terms of learning resources	<ul style="list-style-type: none"> • RLO Composition <p>Formulate:</p> <ul style="list-style-type: none"> • Overview Item (Introduction, Outline, Importance, Pre-requisites) • Content Item • Summary Item (Conclusion, Review, Additional Notes) <ul style="list-style-type: none"> • RIO Composition <p>Formulate:</p> <ul style="list-style-type: none"> • Preview Items • Content Items (Outline, Constraints, Assets Identification and Considerations, Storyboarding) • Practice Items • Assessment Items

2. Instructional Materials Development	Select and develop the instructional materials	̢ Review and select particular learning resources, its pedagogical considerations and instructional strategy ̢ Develop content flowchart ̢ Design graphic user interface and screen templates ̢ Develop storyboarding ̢ Develop scaffolding ̢ Develop interactive multimedia content
--	--	---

Table 3.5 Development Phase

3.4.4 Implementation Phase

Steps	Purposes	Activities
1. Conduct formative evaluation	Review concepts, metaphors, ideas, graphics, and adherent to pedagogical considerations and instructional strategy	̢ Formulate 'beta' version of course ̢ Finalized Course Tree Hierarchy ̢ Finalized Course concept map ̢ Send for evaluation
2. Technical Review	Review all technical issues that may hinder delivery	̢ Determine delivery methods and parameters ̢ Formulate technical specifications (to state the required tools needed for content delivery)

Table 3.6 Implementation Phase

3.4.5 Evaluation Phase

Steps	Purposes	Activities
1. Revision	Evaluate course	̢ Conduct teacher and learner feedbacks ̢ Review learner's performance ̢ Schedule for revisions
2. Conduct summative evaluation	Finalize course delivery matters and synthesize new map with the existing course concept map (based on the theory of learning dependency and knowledge gap)	̢ Synthesize Course concept map ̢ Send for verification
3. Deployment to server	For delivery	̢ Meta-tagged for efficient searches

Table 3.7 Evaluation Phase

3.5 Summary

E-learning materials that have been effectively and carefully designed will facilitate the achievement of desired learning outcomes for the learners. In the absence or reduction of face-to-face communications, the effectiveness of the electronic learning materials will rest on the use of appropriate instructional strategies to mitigate the negative effects of

such losses. It is advocated that such changes in the learning environment require different educational design considerations such as those outlined in this chapter.

Appendices A and B supplement this chapter. Appendix A illustrates the ADDIE Model that is discussed in section 3.3 and 3.4. Based on the Singapore Educational System and following the Ministry of Education (MOE) learning guidelines and teaching curriculum (see <http://www.moe.gov.sg/>), the first two phases of the ADDIE Model, namely the analysis and design phases, are illustrated. Using a science curriculum for secondary two students, all the pedagogical considerations that were outlined in this chapter were discussed. Focusing on the pedagogical nature of learning, the next three phases, namely the development, implementation and evaluation phases, are not discussed since they are typically technical in nature as they look at the provision of content from an implementation and delivery point of view.

Appendix B is used to illustrate the scaffolding framework. Again, only the pedagogical aspects of scaffolding are illustrated. The technical implementations and delivery techniques are presented in chapter six and seven.

CHAPTER 4

LEARNING PERSONALIZATION METHODOLOGY – THEOREMS, ALGORITHMS, AND MATHEMATICAL MODEL

This chapter provides the theoretical background for the learning personalization methodology and consists of three sections. Section 4.1 begins with the formal definitions and explanations of all the concepts underlying the learning personalization methodology. Section 4.2 revisits the learning problems that is presented in chapter 2 to explicitly highlight which learning problems are been addressed by the proposed learning personalization methodology. Section 4.3 discusses the three concept map formats, namely the Cognitive Map (C-Map), Course Expertise Map (E-Map) and Personalized Map (P-Map), and presents the algorithms underlying the formulation of these maps. The documentation of an experimental setup, procedure, results observations and analysis is also presented in this section to justify some of the proposed work. This chapter is summarized in section 4.4. Appendices C and Appendix D supplement this chapter.

4.1 Learning Personalization Concepts

This section presents the concepts underlying our learning personalization approach. The formal definition of learning personalization that is defined in Chapter 1 is reiterated here. Next, our two novel concepts of learning dependency and expertise gap identification are presented. This is followed by the definitions of three concept map formats. These concept maps are used to realize our learning personalization methodology.

4.1.1 Concept – Learning Personalization

Learning personalization is defined as the process of developing individualized learning programs for each learner with the intent of engaging a learning process that is characterized by one's prior knowledge, current needs and learning preferences.

It transforms the content from the content developer's perspective to that of the learner's perspective. Based on the gathering of information about the learner, the learning personalization methodology formats and delivers the right learning content at the right time to the right learner based on the learner's specified learning needs. The learning personalization methodology caters not only to the learning needs or presentation formats of the learning resources, but also takes into account the learning pedagogy considerations that were discussed in Chapter 3.

Learning personalization is a cross-domain study that involves research issues that cross the domain of Information Retrieval, Agent Technologies and Learning Pedagogy. In this research, learning personalization is formulated based on our novel concept of learning dependency and is catered for in two stages: (1) Expertise Gap Identification and (2) Learning Preferences and Constraints Customization.

4.1.2 Concept – Learning Dependency

[Definition 4.1] Learning Dependency is the relation between dependent courses as defined by cognitive- and pedagogical-centered theories governing the mapping of course structure.

This concept is devised from the fundamental principles underlying both cognitivism and constructivism theorems that place most emphasis on the internal mental processes of the

learner's mind and how it can be utilized to promote effective learning. It relies on the explicit elicitation and externalization of the learner's cognitive structure. Such an explicit elicitation of cognitive structure represents the learner's prior knowledge. Through a mapping of the learner's prior knowledge to his targeted knowledge point, a learning map of all the possible learning paths from the learner's current knowledge point to his targeted knowledge point can then be formulated. Learning dependency thus, maps out a learning map that presents all the possible learning routes in a graphical and systematic form. The learning concepts in the learning route are presented in a sequential manner and define the prior and essential learning concepts that must be mastered before certain targeted concepts can be learned.

The concept of learning dependency establishes some form of learning sequence which states that certain pre-requisite concepts must first be mastered before some higher level concepts can be learned. This concept is realized by the application of expertise gap identification that maps out the appropriate learning paths that a learner has to take in order to master a certain desired learning concept.

4.1.3 Concept – Expertise Gap

[Definition 4.2] Expertise Gap is the difference in cognitive expertise between two entities within the same knowledge domain.

Expertise Gap exists when the source and the recipient in a knowledge transfer process possess different levels of expertise. Its existence is a pre-requisite for knowledge transfer whereby knowledge flows from the entity with the higher level of cognitive expertise to the one with the lower level. In our learning context, the entity with the higher cognitive expertise refers to the course while the entity with the lower cognitive expertise refers to the learner.

Figure 4.1 Expertise Gap

Our research goal is therefore to facilitate knowledge transfer. In doing so, this research aims to reduce the cognitive differences between the transfer partners. However, the size

of the cognitive expertise gap between those involved in the process has important implications towards the anticipated success of knowledge transfer. A review of theoretical arguments and empirical evidence reveals that the size of the expertise gap impacts the knowledge transfer process – a small cognitive expertise gap is beneficial for knowledge transfer, whereas too large of a gap actually prohibits transfer [Simon et al., 1974; Ko et al., 2005; Chung et al., 2005; Cho et al., 2006].

4.1.4 Concept – Cognitive Map (C-Map)

[Definition 4.3] Cognitive Map (C-Map) is an externalization of the learner’s learning abilities and expertise in a particular domain. It is an externalization of one’s cognitive structure and prior knowledge.

The C-Map identification is essential to identify the starting point of the Expertise Gap and to align the knowledge transfer according to pedagogical theories. It is based on the learning pedagogy belief that students learn more effectively when the new learning materials are based on (or built upon) (1) a content area where they already know something about, and (2) when the concepts in that area mean something to them and/or to their particular background or culture. When the e-learning content is able to fuse the new information with the student's prior knowledge, it not only aids in the synthesis of the new information into the learner’s cognitive structure, it also triggers his interest and curiosity, and infuses the instruction with a sense of purpose. The prior knowledge thus acts as a lens through which the learner views and absorbs the new information. Refer to Chapter 7, Figure 7.2 for a C-Map example (pp. 191).

4.1.5 Concept – Course Expertise Map (E-Map)

[Definition 4.4] Course Expertise Map (E-Map) is the real time mapping of the essential course curriculum. It represents a visualization of the results of the learner’s query in a concept map format.

The E-Map identification is essential to identify the ending point of the Expertise Gap and to align the knowledge transfer according to pedagogical theories. It is populated dynamically based on the assessment of the learner’s desired learning needs (i.e. through

his query) at each point in time. When a learner wishes to master some learning concept(s), he will need to formulate a query to search for the related concept(s). The query can take the form of either keywords or key phrases and can also be weighted and ranked. The query will be parsed by the search algorithm to ensure that the learner receives the appropriate courses in the shortest possible time. Refer to Chapter 7, Figure 7.3 for an E-Map example (pp. 191).

4.1.6 Concept – Personalized Map (P-Map)

[Definition 4.5] Personalized Map (P-Map) is the personalized learning route(s). It represents the synthesis between the C-Map and E-Map and is a visualization of the possible personalized routes in a concept map format.

P-Map is formed in a 3-step process. First, the possible learning routes are mapped by linking the C-Map (starting point of the expertise gap) to the E-Map (ending point of the expertise gap). Next, the possible learning routes are refined to exclude learning solutions that cannot satisfy the learner's pre-defined learning preferences, constraints and goals. Lastly, the content presentation is personalized according to the learner's learning preferences. Refer to Chapter 7, Figure 7.4 for a P-Map example (pp. 193).

4.2 Learning Personalization Scenario/Problems Revisit

Revisiting the learning personalization scenario in chapter 2, the learning personalization concepts are one of our research solutions that address the existing e-learning content catering problem. Employing the pedagogical content development methodology (in particular, the learning resource structure), the expertise gap identification process is devised to recommend learning solutions that 1) cater to the learner's learning preferences and expertise (through C-Map), and 2) present the training solution as a series of inter-related courses through a visual map (through E-Map and P-Map). The training solution based on our algorithms will be presented in detail in Chapter 7.

The employed algorithm, specifically the assignment of learning weights, is employed to addresses the problem of inaccurate learning solution recommendation due to keyword match technique. An experiment is presented in section 4.3.2.1 to justify our claims and contributions.

4.3 The Algorithms

4.3.1 C-Map

The C-Map is used to store both the academic expertise and everyday experiences (tacit knowledge) of the learner. All newly registered learners will be given an empty C-Map. For ease of searching and synthesis, the C-Map is categorized into the subject domains that are being offered by the system. The C-Map will only be filled upon successful acquisition of knowledge and it can only be populated based on the following methods:

1. (Implicit) Successful completion of a course's post-assessment, or
2. (Explicit) Certified knowledge acquisition (Certificates) by a system administrator;

[Definition 4.6] Domain is the classification of course offerings in a learning context.

Domain classification is essential to allow proper indexing and retrieval of courses. Its application provides the learners / content developers with quick content references, facilitating the search for related courses, as well as searches for work in other online repositories. It also ensures correct placement of courses when a new course is added.

The domain classification in this research adopts the basic classification scheme from the ACM (Association for Computing Machinery) Computing Classification System (CCS) (<http://www.acm.org/class/1998/ccs98.html>). The Nanyang Technological University (NTU) course curriculum classification (<http://www.ntu.edu.sg/publicportal/about+ntu/academics/93dcd4a-f4c3-4356-851b-195149632bac.htm>) is also incorporated.

With respect to the subject area of the research (computing literature, engineering field), CCS is selected as the basis of classification as it has become a standard for identifying and categorizing computing literature, as well as areas of computing interest and/or expertise. However, as CCS does not reflect the academic course categorization, NTU's course curriculum classification is employed as a guide to aid in the classification. The domain classification has two main parts: a numbered tree containing unnumbered subject descriptors, and a General Terms list. The unnumbered subject descriptors are essentially fourth level nodes that involve a four-level tree that has three coded levels and an uncoded level of subject descriptors (usually appearing at the fourth level). This tree comprises the categories and subject descriptors.

4.3.1.1 Expertise Population Rules

When a new learner registers in the system, the system will request the student to take an expertise survey. Based on the expertise survey and our novel expertise population rules, the learner's empty C-Map will be populated. An example of the expertise population rules (mathematics domain) is provided in List 4.1. It is important to note that all the learning concepts and rules are based on the Singapore Educational System and follows the Ministry of Education learning guidelines (see <http://www.moe.gov.sg/>).

Sample of Expertise Population Rules

Domain: Mathematics

Academic Level: Primary

IF pri 1 THEN add concepts

Whole Number I	Mensuration I	Measures I	
Money I	Statistics I	Geometry I	

IF pri 2 THEN add concepts

Whole Number II	Mensuration II	Measures II	
Money II	Statistics II	Geometry II	

IF pri 3 THEN add concepts

Whole Number III	Mensuration III	Geometry III	Fractions I
Money III	Statistics III	Measures III	

IF pri 4 THEN add concepts

Whole Number IV	Mensuration IV	Measures IV	Fractions II
Money IV	Statistics IV	Geometry IV	Decimals I

IF pri 5 THEN add concepts

Whole Number V	Geometry V	Average	Measures V
Money V	Fraction III	Ratio I	Mensuration V
Statistics V	Decimals II	Percentage	Rate

IF pri 5 EM3 THEN add concepts

Whole Number V	Mensuration V	Measures V	Fractions III
Money V	Statistics V	Geometry V	Decimals II

IF pri 6 THEN add concepts

Money VI	Statistics VI	Speed	Algebra
Measures VI	Geometry VI	Ratio II	
Mensuration VI	Time	Proportion	

IF pri 6 EM3 THEN add concepts

Money VI	Statistics VI	Percentage	
Measures VI	Geometry VI	Average	
Mensuration VI	Proportion	Rate	

List 4.1 Expertise Population Rules

After the initial C-Map population, any subsequent modification to the C-Map can only be done through the Cognitive Map Population Algorithm (CMP-Algorithm). Depending on the nature of the population (implicit or explicit), different algorithms are used.

4.3.1.2 CMP-Algorithm 1– CMPA1

CMP-Algorithm 1 – CMPA1

Usage:

CMPA1 is used to populate the C-Map. It is performed implicitly and automatically by the system when the learner successfully completes the course's post-assessment. No form of learner's or SME's input is required.

Input:

$Prof_{expertise}$ – Expertise Profile of User

C_{list} – The course undertaken + its list of pre-requisite courses

Output:

$Prof_{expertise}$ – Updated expertise profile (C-Map)

Methods:

1. **Retrieve** C_{list}
2. **Populate Output Vector** $_{list}$: Capture the expertise and learning preference selections that are associated with each successful course completion. The expertise, the immediate course association (associate, essential pre-requisites, supplementary pre-requisites, augment post-requisites, utilize post-requisites), learning preferences selected, and the completeness index will be synthesized into the output Vector list.

$Output\ Vector_{list}: \langle C_i, Assoc_i, ExpIndex_i, DetIndex_i, PrefIndex_i, Comp_i \rangle$

- C_i represents the course id.
- $Assoc_i$ represents the course immediate association and is formulated as a form of matrix where
 $Assoc_i = \{ C_j, assoc, essentialPreReq, suppPreReq, augPostReq, ultPostReq \}$
 where
 - C_j represents the associated course id
 - $assoc$ represents the association index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is an association of course C_i .
 - $essentialPreReq$ represents the essential pre-requisite course index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is an essential pre-requisite of course C_i .
 - $suppPreReq$ represents the supplementary pre-requisite course index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is a supplementary pre-requisite of course C_i .
 - $augPostReq$ represents the augment post-requisite course index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is an augment post-requisite of course C_i .
 - $ultPostReq$ represents the utilize post-requisite course index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is an utilize post-requisite of course C_i .
- $ExpIndex_i$ represents the expertise index of the course and can only take the values of $\{0, 1\}$. A numerical value of 0 indicates that the course has no specified expertise level while a numerical value of 1 (compulsory for all system-developed courses) indicates that the course is categorized under the system-defined expertise category. An $ExpIndex_i$ of '1' must be associated with a Exp_i value that can only take the values of $\{1, 2, 3\}$ where

$$Exp_i = \begin{cases} 1 & \text{for novice expertise level} \\ 2 & \text{for intermediate expertise level} \\ 3 & \text{for expert expertise level} \end{cases}$$

- $DetIndex_i$ represents the detail index of the course and can only take the values of $\{0, 1\}$. A numerical value of 0 indicates that the course has no specified level of details while a numerical value of 1 (compulsory for all system-developed courses) indicates that the course is categorized under the system-defined level of details category. A $DetIndex_i$ of '1' must be associated with a Det_i value that can only take the values of $\{1, 2, 3\}$ where

$$Det_i = \begin{cases} 1 & \text{for overview} \\ 2 & \text{for normal} \\ 3 & \text{for detailed} \end{cases}$$

- $PrefIndex_i$ represents the preference index of the course and can only take the values of $\{0, 1\}$. A numerical value of 0 indicates that the course has no specified preference level while a numerical value of 1 (compulsory for all system-developed courses) indicates that the course is categorized under the system-defined preference category. A $PrefIndex_i$ of '1' must be associated with a $Pref_i$ value. $Pref_i$ represents the preference index of the learner for a particular course and is formulated as a form of matrix following the Index of Learning Styles (see Chapter 6, section 6.2.4.3.1) where the columns represent the four dimensions of the learning preference. i.e

$$Pref_i = \{\text{active/reflective, sensing/intuitive, visual/verbal, sequential/global}\}$$

where

$$\text{Numerical value of } \begin{cases} 1 & \text{indicates one extreme of the learning preference} \\ 0 & \text{indicates no preference} \\ -1 & \text{indicates the next extreme of the learning preference} \end{cases}$$

For example, a course with a $Pref_i = \{0, -1, 1, 1\}$ states that the course is based on an intuitive, visual and sequential learning preference presentation format.

- $Comp_i$ represents the completeness index and can only take the values of $\{1, 0, -1\}$.

$$Comp_i = \begin{cases} 1 & \text{if the course has been successfully completed} \\ 0 & \text{if the course has not been taken but is associated to the main course} \\ -1 & \text{if the course has been taken but the passing requirements not met} \end{cases}$$

3. **Update Expertise Profile:** Update $Prof_{expertise}$ by synthesizing $Output\ Vector_{list}$ with previous $Prof_{expertise}$ that is recorded in the C-Map.

Algorithm 4.1 CMPA1

4.3.1.3 CMP-Algorithm 2 – CMPA2

CMP-Algorithm 2 – CMPA2

Usage:

CMPA2 is used to populate the C-Map. It is performed explicitly and only upon the learner's request. Upon the execution of the algorithm, the system administrator has to validate the processed results before the output vector is synthesized with the learner's C-Map. This algorithm is necessary and important because (1) the repository's course offerings can never be exhaustive [some certified courses (by other academic boards)

might not be provided by the system] and (2) the learner might take up similar courses offered by other academic institutions. Hence, in order to accurately reflect the learner's existing learning knowledge, the system must take into account all other 'knowledge' that is gained "outside the system". However, the system cannot assume that such 'knowledge' is accurately imparted (or even valid) and hence, the system administrator or the domain expert has to validate the newly acquired knowledge (through validating the certificates or using appropriate assessment strategy to evaluate the learner's knowledge). Once the certification is done, the system administrator must categorize the acquired knowledge into a particular course domain and suggest possible course concept map linkages or associations to the existing system's course offerings (if any).

Input:

$Prof_{expertise}$ – Expertise Profile of User

C_{list} – The course undertaken + its list of pre-requisites courses

Output:

$Prof_{expertise}$ – Updated expertise profile

Methods:

1. **Initialize input nodes, output nodes, connection weights:** Assign the main course name and its pre-requisite courses names (up to a maximum of 3 levels) as individual nodes in a two dimensions (2-D matrix) input vector. Initialize node coordinate to 0; the node coordinate is used to determine whether the proposed course name corresponds to the system's predefined course name.

Input Vector: $\langle \text{Course}, n_i \rangle$

Using the N terms input vector, set up the connection vector by creating a three-dimensional map (3-D matrix) of N output nodes (N by 3 Matrix): post-requisite course C_i , pre-requisite course C_j , and hierarchical level h_{ij}^p . The hierarchical level h_{ij}^p represents the difference in the hierarchical level between C_i and C_j . Note that the superscript p states that these levels are proposed and is pending verification by the SMEs. C_i represents the main course while C_j represents existing pre-requisite course to the main course. All h_{ij}^p are initialized to the learner's certificate proposed hierarchical level.

Output Vector_{temp}: $\langle C_i, C_j, h_{ij}^p \rangle$

2. **Present document in ontological names:** Using the domain-specific ontology, describe each document (input and output vector) using the pre-defined ontological names. Set each node coordinate, n_i , in the input vector to 1 if the course name has a corresponding ontological term that is recognized by the system. Each document is presented to several ontologies. Should a certain node coordinate remain at 0 after several attempts, the system administrator will be notified.

Note: For each domain, the SMEs will decide (1) how many failed attempts, (2) the number of different ontologies to be employed and (3) whether to include cross-domain searches before the system administrator is notified.

Output Vector_{final}: $\langle \text{Course}, \text{baseName}, C_i \rangle$

where C_i represents the course id (automatically assigned by the system)

3. **Populate Output Vector_{list}:** Capture the expertise and learning preference selections that is associated with each course. The selected expertise (if any), the immediate course association (associate, essential pre-requisites, supplementary pre-requisites,

augment post-requisites, utilize post-requisites), the selected learning preferences (if any), and the completeness index will be synthesized into the output Vector list.

Output Vector_{list}: $\langle C_i, Assoc_i, ExpIndex_i, DetIndex_i, PrefIndex_i, Comp_i \rangle$

- C_i represents the course id.
- $Assoc_i$ represents the course immediate association and is formulated as a form of matrix where
 $Assoc_i = \{ C_j, assoc, essentialPreReq, suppPreReq, augPostReq, ultPostReq \}$
 where
 - C_j represents the associated course id
 - $assoc$ represents the association index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is an association of course C_i .
 - $essentialPreReq$ represents the essential pre-requisite course index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is an essential pre-requisite of course C_i .
 - $suppPreReq$ represents the supplementary pre-requisite course index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is a supplementary pre-requisite of course C_i .
 - $augPostReq$ represents the augment post-requisite course index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is an augment post-requisite of course C_i .
 - $ultPostReq$ represents the utilize post-requisite course index and can only take the value of $\{0,1\}$ with 1 indicating that course C_j is an utilize post-requisite of course C_i .
- $ExpIndex_i$ represents the expertise index of the course and can only take the values of $\{0, 1\}$. A numerical value of 0 indicates that the course has no specified expertise level while a numerical value of 1 (compulsory for all system-developed courses) indicates that the course is categorized under the system-defined expertise category. An $ExpIndex_i$ of '1' must be associated with a Exp_i value that can only take the values of $\{1, 2, 3\}$ where

$$Exp_i = \begin{cases} 1 & \text{for novice expertise level} \\ 2 & \text{for intermediate expertise level} \\ 3 & \text{for expert expertise level} \end{cases}$$
- $DetIndex_i$ represents the detail index of the course and can only take the values of $\{0, 1\}$. A numerical value of 0 indicates that the course has no specified level of details while a numerical value of 1 (compulsory for all system-developed courses) indicates that the course is categorized under the system-defined level of details category. A $DetIndex_i$ of '1' must be associated with a Det_i value that can only take the values of $\{1, 2, 3\}$ where

$$Det_i = \begin{cases} 1 & \text{for overview} \\ 2 & \text{for normal} \\ 3 & \text{for detailed} \end{cases}$$
- $PrefIndex_i$ represents the preference index of the course and can only take the values of $\{0, 1\}$. A numerical value of 0 indicates that the course has no specified preference level while a numerical value of 1 (compulsory for all system-developed courses) indicates that the course can be categorized under the system-defined preference category. A $PrefIndex_i$ of '1' must be associated with a $Pref_i$

value. $Pref_i$ represents the preference index of the learner for a particular course and is formulated as a form of matrix following the Index of Learning Styles where the columns represent the four dimensions of the learning preference. i.e.

$Pref_i = \{\text{active/reflective, sensing/intuitive, visual/verbal, sequential/global}\}$
where

Numerical value of $\begin{cases} 1 \text{ indicates one extreme of the learning preference} \\ 0 \text{ indicates no preference} \\ -1 \text{ indicates the next extreme of the learning preference} \end{cases}$

For example, a course with a $Pref_i = \{0, -1, 1, 1\}$ states that the course is based on an intuitive, visual and sequential learning preference presentation format.

- $Comp_i$ represents the completeness index and can only take the values of $\{1, 0, -1\}$.

$Comp_i = \begin{cases} 1 \text{ if the course has been successfully completed} \\ 0 \text{ if the course has not been taken but is associated to the main course} \\ -1 \text{ if the course has been taken but the passing requirements not met} \end{cases}$

4. **Update Expertise Profile:** Update $Prof_{expertise}$ by synthesizing $Output\ Vector_{list}$ with previous $Prof_{expertise}$.

Algorithm 4.2 CMPA2

4.3.2 E-Map

The E-Map relies on the usage of an efficient search algorithm to determine the type and amount of course documents to retrieve. The algorithm is based on the concept of relevancy and makes use of the basic concept of the Vector Model. The basic concept underlying the Vector Model is presented in Appendix C.

The main theory adopted from the Vector Model is the cosine similarity measure. The cosine similarity measure is a popular measure of similarity for text (which normalizes the features by the covariance matrix) clustering. The cosine measure is the cosine of the angle between two vectors and is given by:

$$sim(d_i, q) = \frac{\vec{d_i} \cdot \vec{q}}{\left| \vec{d_i} \right| \times \left| \vec{q} \right|} = \frac{\sum_{i=1}^t w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^t w_{i,j}^2} \times \sqrt{\sum_{i=1}^t w_{i,q}^2}} \quad (4.1)$$

Eq. 4.1 captures a scale invariant understanding of similarity. An even stronger property is that the cosine similarity does not depend on the length. Hence, this allows documents with the same composition, but different totals to be treated identically which makes this

the most popular measure for text documents. Also, due to this property, samples can be normalized to the unit sphere for more efficient processing [Dhillon, 2001].

In this research, we use the cosine measure to represent the similarity between the query and the course documents. We let:

$$\text{Document, } \vec{d_j} = (w_{1j}, w_{2j}, \dots, w_{ij}) \text{ and Query, } \vec{q} = (w_{1q}, w_{2q}, \dots, w_{iq}).$$

In order to apply the $\text{sim}(d_j, q)$ value as a form of ranking mechanism for our learning context, an important parameter, the index term weightings, must first be obtained. The index term weights computation has been used to indicate how important a term is for a particular document. There are many ways in which the index term weights can be computed. For example, one of the most popular term weighting schemes assigns high weighting to terms that occur frequently in a given document (intra-cluster similarity) but infrequent in the rest of the collection (inter-cluster dissimilarity). This method works on the principle that such terms are good discriminators for a document and are termed formally as the clustering analogy. In the clustering terminology, intra-cluster similarity is used to assign importance to the term that appears often in a document while inter-cluster dissimilarity is used to assign importance to a term's scarcity across the collection of documents (refer to Appendix C for the basic principle for the clustering analogy).

However, the unique characteristics of our learning environment (i.e. the distributed yet well categorized and tagged nature of the distance learning materials, the need for personalized learning routes formulation, etc.) requires adjustment in the nature of its information retrieval techniques specifically designed for that modality. Hence, while the widely employed tf-idf weighting scheme is appropriate for many IR applications, its usage in our learning environment is rather inappropriate. This is because the learning environment is concerned not only with the issue of information retrieval, but also strives to provide personalized learning routes that are catered dynamically according to the learner's learning expertise and preferences. Thus, in order to cater to our proposed higher-level cognitive strategies and to take into account the hierarchical nature of the course documents, a personalized and unique search and retrieval technique has to be devised.

The unique nature of the learning resources is worth noting.

- All documents in the learning repository are course materials;
- The course materials can reside at different levels of hierarchy (i.e. RLO-level, RIO-level, etc.) or domains;
- Each course document in the learning repository is described by a standardized set of metadata;
- A sub-category, that states the key learning concepts that will be covered by the course, is listed within the set of course metadata.

Unlike IR on Web documents where the index terms are based on the frequency of keyword occurrences, in a learning domain, the assignment of index terms by keyword occurrences is not appropriate. This is because in most cases, a particular course that teaches some key concepts might not even contain the keyword itself in the learning materials. Also, in most cases, the important key concepts are usually not the term that occurs the most frequently in the course materials. Instead, it is proposed that the key learning concepts (which reside in the course metadata) be set as the index terms for the course document.

In order to prove this point, an experiment is carried out using 4 post graduate engineering courses that were conducted in the Nanyang Technological University of Singapore. From these 4 courses, a total of 42 modules are selected for the experiment.

4.3.2.1 Experimental Setup

This experiment is carried out to investigate the appropriateness of assigning index terms and hence their weightings for learning courses based on keyword occurrences. Current Web search engines such as AltaVista (www.altavista.com) and Google (www.google.com) typically treat the natural language query as a list of terms and retrieve documents that are similar to the original query. Similarity is gauged based on the query terms occurrence in the documents and documents that contain the highest frequency of the query terms' occurrences will be ranked the highest. However, it is argued that the assignment of index terms based on query keywords or key phrases occurrence is inappropriate for the learning environment. Learning materials with the best answers may

contain few of the terms from the original query and ranked low by the learning environment search engine. Also, such a flat representation does not lend itself to the hierarchical learning materials structure and cannot fully reflect the learner's learning preferences in terms of the weights assignment.

In order to prove our stand, a test sample consisting of 4 engineering post graduate courses that were conducted in the Nanyang Technological University of Singapore is selected. These selected courses are (1) software engineering, (2) system engineering (and graph theory), (3) networking and database and (4) knowledge engineering. In order for accurate analysis, a total of $(11 + 9 + 9 + 13) = 42$ modules in the 4 courses were used.

4.3.2.1.1 Experimental Procedure

In this section, the experiment procedure is discussed. The main governing principle behind the experiment is as follows:

Given a course document, d_i :

1. The document is parsed to remove common stopwords
2. The document is parsed to remove domain specific stopwords
3. The document is stemmed to reduce the word to its stem or root form
4. The top 10 most frequent occurring term is set as the $Vector_{freq}$ where

$$Vector_{freq} = \left\{ \begin{array}{l} \text{Document Term}_1, w_{1,f} \\ \text{Document Term}_2, w_{2,f} \\ \vdots \\ \text{Document Term}_{10}, w_{10,f} \end{array} \right\}$$

and

$$w_{i,f} = \frac{rank}{w_{\max}} = \frac{rank}{\sum_{j=1}^{10} j} \quad (4.2)$$

i.e. the most frequently occurring term will be hold the numerical rank of 10.

Hence, it's corresponding $w_{1,f} = \frac{rank}{w_{\max}} = \frac{10}{55} = 0.182$. The next frequent occurring term will hold a numerical rank that is 1 less than the previous one (i.e. 9) and so on and so fro.

5. The list of key learning concepts is populated and stored as the $Vector_{keyCon}$

$$\text{Vector}_{\text{keyCon}} = \left\{ \begin{array}{l} \text{Concept Term}_1, w_{1,k} \\ \text{Concept Term}_2, w_{2,k} \\ \dots\dots\dots \\ \text{Concept Term}_{10}, w_{10,k} \end{array} \right\}$$

and $w_{i,k}$ is the weightings assigned by the SMEs.

The key concepts are also assigned by the SMEs. However, for testing purposes, the weightings are normalized to be a constant average value. Therefore,

$$w_{i,k} = \frac{1}{\text{NoOfKeyConcepts}}$$

i.e. for a total of 5 key learning concepts, all the key concepts will have the same

$$\text{weighting of } w_{i,k} = \frac{1}{\text{NoOfKeyConcepts}} = \frac{1}{5} = 0.2$$

6. The similarity between $\text{Vector}_{\text{freq}}$ and $\text{Vector}_{\text{keyCon}}$ will be calculated using $\text{sim}(\text{Vector}_{\text{freq}}, \text{Vector}_{\text{keyCon}})$

4.3.2.1.2 Experimental Key Processes Explained

1. Parsing of Common Stopword

The search query is parsed through the usage of a common stopwords list. Stopwords are words that may be entered into a query statement but are removed as individual words and not searched. These words occur frequently in documents and they do not carry any specific meaning (such as ‘the’). As such words rarely relate to the meaning or value of a document, they are skipped and will not be in the index nor searchable. Thus, these words are often removed from a search query to simplify the size of the indexing structure. However, in special cases whereby the learner specifies a key phrase query, stopwords parsing will not be carried out as the exclusion of common words can be problematic precisely because they are common and appear often in phrases. Therefore, in normal searches, the learning environment search engine will index everything, and then filter on stopwords and minimum term length at query time unless the query is phrase-based.

A list of common stopwords is compiled and a sample list is shown in Table 4.1.

A	because	from	just	onto	the	when
about	been	front	last	or	their	where
above	become	get	least	our	them	who
after	before	getting	left	ourselves	then	who’s

again	being	go	less	out	there	whoever
Ago	between	goes	let	over	these	whom
All	but	going	like	per	they	whose
almost	by	gone	make	perhaps	this	why
alone	came	got	many	put	those	will
Also	can	gotten	may	putting	through	with
always	cannot	ha	maybe	same	thru	within
Am	com	had	me	saw	till	without
An	come	has	mill	see	to	won't
And	could	have	mine	seen	too	would
another	did	having	more	shall	two	wouldn't
Any	do	he	most	she	unless	yet
anybody	does	her	much	should	until	you
anyhow	doing	here	my	so	up	you've
anyone	done	him	myself	some	upon	your
anything	down	his	never	somebody	us	
anyway	each	how	no	someone	very	
Are	else	i	none	something	was	
As	even	if	not	stand	we	
At	ever	in	now	such	went	
away	every	into	of	sure	were	
back	everyone	is	off	take	what	
Be	everything	isn't	on	than	whatever	
became	for	it	one	that	what's	

Table 4.1 Section of common stopwords used in the experiment

2. Parsing of Domain Specific Stopwords

Stopword lists are not only language specific but also domain specific. For example, a corpus of chemical descriptions would likely benefit from a different stopwords list than a corpus of engineering scientific terms would. Hence, after the search query is parsed by the common stopwords list, it will go through another phrase of parsing by a domain specific stopwords list. This domain specific list is essential to further reduce the query's term size, speed up the processing and to narrow down the scope of the query. Some examples of domain-specific stopwords are provided in Table 4.2.

Domain	Stopwords			
Politics	government	state	city	society
Medicine	patient	doctor	health	healthcare
Education	learner	instructor	student	teacher

Table 4.2 Section of domain specific stopwords used in the experiment

3. Synonyms and stemming application

After the process of removing stopwords (common and domain specific), the search engine will perform stemming on the remaining terms. That is, to reduce all the index terms to their root form. Most languages have closely related words, known as morphological or inflectional variants, based on a common portion known as a stem, or root word. For example, word like ‘running’, ‘runs’, ‘ran’ and ‘runner’ will be transformed to the stem form of ‘run’. Note that stemming can also be applied to word with irregular endings (such as from ‘ran’ to ‘run’ even though the stem ‘run’ does not appear in the variant). Similar to the removal of stopwords, stemming not only increases the likelihood of a successful search, it also decreases the index size. After stemming, the list of key words will go through a synonyms transformation to improve the search quality. For example, a query term of ‘car’ will also include searches that use ‘automobile and ‘vehicle’. A ‘substitution list’ will be used to associate alternate terms as the keyword’s synonyms.

4. $\text{Vector}_{\text{freq}}$ population

The top 10 most frequent occurring term is set as the $\text{Vector}_{\text{freq}}$ where

$$\text{Vector}_{\text{freq}} = \left\{ \begin{array}{l} \text{Document Term}_1, 0.182 \\ \text{Document Term}_2, 0.164 \\ \text{Document Term}_3, 0.145 \\ \text{Document Term}_4, 0.127 \\ \text{Document Term}_5, 0.109 \\ \text{Document Term}_6, 0.091 \\ \text{Document Term}_7, 0.073 \\ \text{Document Term}_8, 0.055 \\ \text{Document Term}_9, 0.036 \\ \text{Document Term}_{10}, 0.018 \end{array} \right\}$$

5. $\text{Vector}_{\text{keyCon}}$ population

$$\text{Vector}_{\text{keyCon}} = \left\{ \begin{array}{l} \text{Concept Term}_1, \frac{1}{\sum \text{KeyConcepts}} \\ \text{Concept Term}_2, \frac{1}{\sum \text{KeyConcepts}} \\ \vdots \\ \text{Concept Term}_n, \frac{1}{\sum \text{KeyConcepts}} \end{array} \right\}$$

6. $\text{sim}(\text{Vector}_{\text{freq}}, \text{Vector}_{\text{keyCon}})$ calculation

Applying equation (4.1),

$$\text{sim}(d_i, q) = \frac{\vec{d}_i \cdot \vec{q}}{\left| \vec{d}_i \right| \times \left| \vec{q} \right|} = \frac{\sum_{i=1}^t \text{Vector}_{\text{KeyCon}} \times \text{Vector}_{\text{freq}}}{\sqrt{\sum_{i=1}^t \text{Vector}_{\text{KeyCon}}^2} \times \sqrt{\sum_{i=1}^t \text{Vector}_{\text{freq}}^2}} \quad (4.3)$$

4.3.2.1.3 Experimental Assumptions/Simplifications

For experimental accuracy and simplification purposes, the most frequent document terms will only be ranked according to a single keyword. Key phrase (2 or more words) is excluded. This is because in the actual searches, the search query (keywords or key phrases) is already formulated. However, for this experiment, we are testing the appropriateness of assigning index terms and hence their weightings for learning courses based on keyword occurrences. Therefore, it is not appropriate or even possible to formulate logical key phrase to be included in the frequency ranking. Hence, only single keywords will be listed under the most frequent document terms. However, when a particular concept term is made up of a key phrase, that particular key phrase will be indexed to check if its frequency of occurrences falls under the top ten most frequent occurring terms. Should that phrase occur more frequently than any of the top ten (single keyword) terms, that phrase will be inserted into the list. Also, if the concept term has the same frequency as two or more terms in the list of frequent terms, the concept term will be set as the one with the highest ranking. This is because this experiment is constructed with the intention to prove that frequency is not a good gauge to serve as an index. Therefore, when a concept term occurs as one of the more frequently occurring terms, we assign it the highest possible weightage so as to maintain a higher tolerance level.

4.3.2.1.4 Experimental Data

No.	Course	Module
01	Software Engineering	An Introduction to Software Architecture
02		Overview of Software Engineering 1
03		Overview of Software Engineering 2
04		Overview of Software Engineering 3
05		Object-Oriented Concepts and Notations I
06		Object-Oriented Concepts and Notations II
07		Object-Oriented Software Development Process
08		Some Generic Object-Oriented Design
09		Object-Oriented Analysis and Design Using UML - Case Studies
10		Measuring Object-Oriented Design

Chapter 4: Learning Personalization Methodology – Theorems, Algorithms and Mathematical Model

11		Emerging Topics in Software Engineering
12	System Engineering (and Graph Theory)	Design of System Architectures
13		Systems Engineering
14		System Concepts
15		Sociotechnical Systems
16		Using the WinWin Spiral Model: A Case Study
17		Introduction to Graph Theory
18		Graphs- Basics
19		Graph Applications I
20		Graph Applications II
21	Networking and Databases	Network Design: ISO/OSI reference model
22		Network Design: Local Area Networks
23		Network Design: Internet TCP/IP and Internetworking
24		Network Design: Industrial Networking - Industrial LAN
25		Network Design: Internet Application and Issues
26		Network Design: Web Application Development Tools
27		Network Design: Industrial Networking - Serial Networking
28		Database and Middleware: Database and Middleware Introduction
29		Database and Middleware: Entity-Relationship models
30	Knowledge Engineering	Knowledge Engineering Overview I
31		Knowledge Engineering Overview II
32		Agent Introduction I: Why, When, and Where to Use Software Agents
33		Agent Introduction II: Agent-Based Systems for Intelligent Manufacturing: A State-of-the-Art Survey
34		Concurrent Knowledge Processing Techniques I
35		Concurrent Knowledge Processing Techniques II
36		Autonomous Agents/Actors I
37		Autonomous Agents/Actors II
38		Machine Learning
39		Evolutionary Computation
40		Data Mining and Knowledge Discovery
41		Autonomous agent - AI Planning
42		Sub-Symbolic Learning

Table 4.3 Experimental Data**4.3.2.1.5 Experimental Sample**

A sample of one of the courses, An Introduction to Software Architecture, is used as an illustration in Table 4.4. The similarity between $Vector_{freq}$ and $Vector_{keyCon}$ is calculated using $\text{sim}(Vector_{freq}, Vector_{keyCon})$ and illustrated below. Please refer to Appendix D for the full experiment data and its calculation.

No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	architecture	165	system	198
02	design	65	architecture	165
03	architectural styles	24	data	153
04	case studies	20	software	151
05			component	75
06			control	73
07			model	69
08			level	68

09		design	65
10		blackboard	50

Table 4.4 Course Sample: An Introduction to Software Architecture

The concept terms list the key learning concepts that are taught for a particular course document. These concept terms are listed by the SMEs as the key learning concepts or teaching objectives of the course.

$$\begin{aligned}
 \text{Vector}_{\text{freq}} &= \left\{ \begin{array}{l} \text{system} \quad , \quad 0.182 \\ \text{architecture} \quad , \quad 0.164 \\ \text{data} \quad , \quad 0.145 \\ \text{software} \quad , \quad 0.127 \\ \text{component} \quad , \quad 0.109 \\ \text{control} \quad , \quad 0.091 \\ \text{model} \quad , \quad 0.073 \\ \text{level} \quad , \quad 0.055 \\ \text{design} \quad , \quad 0.036 \\ \text{blackboard} \quad , \quad 0.018 \end{array} \right\} \\
 \text{Vector}_{\text{KeyCon}} &= \left\{ \begin{array}{l} 0 \quad , \quad 0 \\ 1 \quad , \quad 0.25 \\ 0 \quad , \quad 0 \\ 0 \quad , \quad 0 \\ 0 \quad , \quad 0 \\ 0 \quad , \quad 0 \\ 0 \quad , \quad 0 \\ 0 \quad , \quad 0 \\ 1 \quad , \quad 0.25 \\ 0 \quad , \quad 0 \end{array} \right\} \\
 \text{sim}(\text{Vector}_{\text{freq}}, \text{Vector}_{\text{KeyCon}}) &= \frac{\vec{d}_i \cdot \vec{q}}{\left| \vec{d}_i \right| \times \left| \vec{q} \right|} = \frac{\sum_{i=1}^t \text{Vector}_{\text{KeyCon}} \times \text{Vector}_{\text{freq}}}{\sqrt{\sum_{i=1}^t \text{Vector}_{\text{KeyCon}}^2} \times \sqrt{\sum_{i=1}^t \text{Vector}_{\text{freq}}^2}} \\
 &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.25, 0, 0, 0, 0, 0, 0, 0.25, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\
 &= 0.396
 \end{aligned}$$

4.3.2.1.6 Experimental Results

No.	Module	sim(Vector _{freq} , Vector _{KeyCon})
01	An Introduction to Software Architecture	0.396
02	Overview of Software Engineering 1	0.230
03	Overview of Software Engineering 2	0.238
04	Overview of Software Engineering 3	0.432
05	Object-Oriented Concepts and Notations I	0.490
06	Object-Oriented Concepts and Notations II	0.260

Chapter 4: Learning Personalization Methodology – Theorems, Algorithms and Mathematical Model

07	Object-Oriented Software Development Process	0.325
08	Some Generic Object-Oriented Design	0.144
09	Object-Oriented Analysis and Design Using UML - Case Studies	0.401
10	Measuring Object-Oriented Design	0.191
11	Emerging Topics in Software Engineering	0.325
12	Design of System Architectures	0.650
13	Systems Engineering	0.144
14	System Concepts	0.260
15	Sociotechnical Systems	0.024
16	Using the WinWin Spiral Model: A Case Study	0.201
17	Introduction to Graph Theory	0.410
18	Graphs- Basics	0.000
19	Graph Applications I	0.000
20	Graph Applications II	0.287
21	Network Design: ISO/OSI reference model	0.431
22	Network Design: Local Area Networks	0.606
23	Network Design: Internet TCP/IP and Internetworking	0.403
24	Network Design: Industrial Networking - Industrial LAN	0.107
25	Network Design: Internet Application and Issues	0.376
26	Network Design: Web Application Development Tools	0.404
27	Network Design: Industrial Networking - Serial Networking	0.047
28	Database and Middleware: Database and Middleware Introduction	0.577
29	Database and Middleware: Entity-Relationship models	0.795
30	Knowledge Engineering Overview I	0.346
31	Knowledge Engineering Overview II	0.289
32	Agent Introduction I: Why, When, and Where to Use Software Agents	0.577
33	Agent Introduction II: Agent-Based Systems for Intelligent Manufacturing: A State-of-the-Art Survey	0.289
34	Concurrent Knowledge Processing Techniques I	0.404
35	Concurrent Knowledge Processing Techniques II	0.116
36	Autonomous Agents/Actors I	0.000
37	Autonomous Agents/Actors II	0.000
38	Machine Learning	0.000
39	Evolutionary Computation	0.317
40	Data Mining and Knowledge Discovery	0.336
41	Autonomous agent - AI Planning	0.217
42	Sub-Symbolic Learning	0.000

Table 4.5 $\text{sim}(\text{Vector}_{\text{freq}}, \text{Vector}_{\text{KeyCon}})$ Value

4.3.2.1.7 Experimental Results Analysis

Sim() Range	No. of Modules	Percentile
0.0-0.1	8	19
0.1 ⁺ -0.2	5	12
0.2 ⁺ -0.3	9	21
0.3 ⁺ -0.4	7	17
0.4 ⁺ -0.5	8	19
0.5 ⁺ -0.6	2	5
0.6 ⁺ -0.7	2	5
0.7 ⁺ -0.8	1	2
0.8 ⁺ -0.9	0	0
0.9 ⁺ -1.0	0	0

Table 4.6 Experimental Result Percentile

Observations

From Table 4.6, it is seen that the majority (about 21%) of the $\text{sim}(\text{Vector}_{\text{freq}}, \text{Vector}_{\text{KeyCon}})$ value falls within the range of $0.2^+-0.3$. The interesting fact to note is that only 12% of the $\text{sim}(\text{Vector}_{\text{freq}}, \text{Vector}_{\text{KeyCon}})$ values are above the mid-value of 0.5 and a high percencile (19%) of the result falls in the 0.0-0.1 range with 6 modules recording a $\text{sim}(\text{Vector}_{\text{freq}}, \text{Vector}_{\text{KeyCon}})$ value of 0.0.

Analysis

It is shown by the experimental results that when the documents are housed in a certain context (i.e. domain of discourse) with some distinct features (such as indexable and trustable metadata), the assigning of index terms based purely on keywords occurrences is not sufficient. The experiment has shown that out of 42 modules, only 5 modules have a similarity index that is above the mid-point value of 0.5 and only one module records a 79.5% match (highest ranking). The fact that 8 modules fall under the 0.0-0.1 range is worth taking note. What is more alarming is the total mismatch (0%) of 6 modules.

Hence, in view of the low similarity values, this experiment concludes that it is inappropriate to assign index terms based purely on the keyword occurrences. While Web documents are still relying on using the keyword occurrences indexing, the main reason underlying its usage is because it is difficult, if not impossible, to index every single document according to a pre-defined standard. Also, although the use of metadata has somehow eased the problem, the inaccurate and inappropriate use of metadata (either due to ignorance or unscrupulous attempt to increase citation) has been well documented [Fritch, 2003; Haase, 2004].

Besides, keyword searches have its inherent problems. Keyword searches cannot distinguish between words that are spelled in the same way but have different semantics (i.e. set theory, a set of utensils, and a set of a badminton game). This often results in hits that are completely irrelevant to the query. Also, the use of keyword searches cause problem with the stem form. For example, if the search query contains the word ‘test’, should the search engine returns a hit on words such as ‘testing’ or ‘tester’? What about the usage of singular and plural form? What about verb tenses that differ from the query

word by only a ‘s’, ‘t’ or ‘ed’? Besides, the search engine cannot return hits on keywords that contain the same semantics but are entered differently from the search query (i.e. a query will not return a document that employs the word ‘cardiac’ instead of ‘heart’).

Therefore, in view of the abovementioned explanations and experimental results, the system employs the use of key learning concepts as the index terms for course documents.

4.3.2.2 Hierarchical Weightage Considerations

While the selection of key learning concepts as the course document’s index terms have been justified, the nature of the course documents (i.e. the granularity and the hierarchical structuring of the courses and the principle basis of learning) require some modifications in the assigning of weightage for each learning concept.

In most learning instances, the learning concepts or the knowledge that the learner wishes to master usually does not reside in a single course or at a particular course hierarchical level; i.e. in order to master all the learning concepts and without the consideration of the learner’s existing learning abilities, the learner might have to take more than one course and these courses might reside at different hierarchical levels (e.g. Topic and Module level) or may even be cross-domain. For example, given a learner’s query of “the concept of Routh-Hurwitz Stability Criterion, Relative Stability and Stability of State Variable Systems”, these concepts while covered under the broad umbrella of the concept of stability, are presented in three courses which reside at different hierarchical levels – At the Module level: Relative Stability and Stability of State Variable Systems; At the Topic level: concept of Routh-Hurwitz Stability Criterion. Thus, depending on the granularity / details of the required learning concepts, not all the key concepts could be presented at the course-level or reside at a single course; that is, the course level will generally teach some higher level concepts while the lower granularity level will teach the more detailed concepts and is typically related to a more specific application context. Therefore, while a particular course may teach certain concepts, these concepts might not be reflected in the highest modularity level. Hence, a weighting algorithm will be investigated to take into account the lower level concepts as it is only reasonable that the further the hierarchal

level, the lower the weightings that will be assigned to that concept to prevent a diversified learning solution.

Thus, in order to provide a complete learning experience that aims to cover all aspects of the learner's requirements, the system may need to present the training solution that is based on more than one course. While the courses exist on a stand-alone basis in the course repository, the system will explicitly elicit the relationship between these courses to bridge the gap between the conceptual and procedural knowledge [Khan, 1993]. Thus, instead of presenting the courses as independent learning materials, the training solution will consist of a series of inter-related courses (whenever it is necessary). The visual image of the learning paths is presented through the usage of the concept mapping technique and labeled as the E-Map.

The E-Map will assign priority weightings in a normalized form according to the ranking algorithm depicted in equation 4.4.

$$rank = \sum_{i=1}^n (found_i \times weighting_i \times \Delta domain_i \times \Delta expertise_i \times \Delta detail_i) \times \frac{1}{\sum course \times \sum level} \quad (4.4)$$

where:

- $found_i$ is a boolean value (either 0 or 1) representing whether a particular search query term has been successfully found. A numerical value of 1 indicates that the search query term is found, where a numerical value of 0 indicates that the search query term is not found.
- $weighting_i$ represents the weightage that is assigned to a successful retrieval of a particular search query term. The value will range between $\{0,1\}$. This value is different from the previous weighting calculation as it takes both the learner's and the content expert's perspective of the query term-document in consideration and is given by:

$$weighting_i = Lweighting_i \times Eweighting_i$$

where

- $Lweighting_i$ represents the importance of satisfying (or finding) a particular query term in the learner's perspective

In the event where the learner does not specify the value, the weightings will be normalized according to the total number of query terms, i.e.

$$Lweighting_i = \frac{1}{NoOfQueryTerms}$$

- $Eweighting_i$ represents the importance of a particular query term to the overall objective of the course document. It is defined from the content expert's perspective.

It is important to take both the $Lweighting_i$ and $Eweighting_i$ value into consideration as the significance of the successfully finding and retrieval of a particular query term will vary with the nature of the course documents. For example, given the query of 'java' in a programming context, a total of 2 modules provide the learning materials (Appendix D). The first module, Object-Oriented Concepts and Notations I, covered by the Software Engineering Course, is designed solely for java programming and teaches some core programming concepts of association, inheritance, generalization, aggregation and constraint. However, the second module, Network Design: Web Application Development Tools, covered by the Networking and Database Course, only introduces Java programming and has the bulk of the learning materials teaching other concepts like Internet, Web, Web Services and Data Presentation. Hence, given the same query term occurrence but different emphasis, it is imperative to consider both the learner's and the content expert's perspective on the relative weightage of the index term – query term occurrence.

The $weighting_i$ values are constrained by equation 4.5.

$$\sum weighting_i = 1 \quad (4.5)$$

- $\Delta domain_i$ represents a difference between the retrieved concept's domain and the query's specified domain. $\Delta domain_i$ can take the value of either $\{1, 0.5\}$. $\Delta domain_i$ will be assigned a value of 1 when there is no change in the retrieved course and learner's selected domain of interest. $\Delta domain_i$ will be assigned a value of 0.5 when the retrieved course is of a different domain from the specified query domain. However, in the event that the learner specifies that no change of domain is tolerated, then $\Delta domain_i$ will take the form of a boolean value (either 0

or 1) where a numerical value of 1 represents no change in the concept domain while a numerical value of 0 represents that a concept has matched the query keyword but belongs to a different domain. While the system does not eliminate cross-domain courses from the search, a successful matching of a cross-domain query-concept term will not add to the ranking of the term. Nonetheless, the concept will still be retrieved but is highlighted in another color in the concept map to differentiate the cross-domain context.

- $\Delta\text{expertise}_i$ represents a difference between the retrieved concept's expertise and the query's specified expertise. $\Delta\text{expertise}_i$ can take the value of $\{1, 0.75, 0.5\}$. A numerical value of 1 indicates no change in the expertise while a numerical value of 0.75 indicates a one level change of expertise and a numerical value of 0.5 indicates a two level change of expertise. Similarly, in the event when the learner specifies that no change of expertise is tolerated, then $\Delta\text{expertise}_i$ will take the form of a boolean value (either 0 or 1) where a numerical value of 1 represents no change in the concept expertise while a numerical value of 0 represents that a concept has matched the query keyword but belongs to a different expertise level. While the system does not eliminate cross-expertise courses from the search, a successful matching of such a query-concept term will not add to the ranking of the term. Nonetheless, the concept will still be retrieved but is highlighted in another color in the concept map to differentiate the cross-expertise context.
- Δdetail_i represents a difference between the retrieved concept's level of detail and the query's specified level of detail. Δdetail_i can take the value of $\{1, 0.75, 0.5\}$. A numerical value of 1 indicates no change in the level of details while a numerical value of 0.75 indicates a one level change of detail level and a numerical value of 0.5 indicates a two level change of detail level. Similarly, in the event when the learner specifies that no change of detail level is tolerated, then Δdetail_i will take the form of a boolean value (either 0 or 1) where a numerical value of 1 represents no change in the concept expertise while a numerical value of 0 represents that a concept has matched the query keyword but belongs to a different expertise level. While the system does not eliminate cross-detail courses from the search, a successful matching of such a query-concept

term will not add to the ranking of the term. Nonetheless, the concept will still be retrieved but is highlighted in another color in the concept map to differentiate the cross-detail context.

- $\sum\text{level}$ represents the total number of different levels that spans the whole learning path. This value can be determined by the learner and is restricted by the system to a maximum of 5 levels to minimize the search time. A diversified learning path is also prevented through the usage of the $\sum\text{level}$ factor.
- $\sum\text{course}$ represents the total number of courses that spans the whole learning path. This value can be determined by the learner. This value will not be restricted by the system as each additional course represents an enhancement to the learning experience. However, a search result with a high number of courses will be penalized in the form of a lower ranking.

$\sum\text{level}$ and $\sum\text{course}$ are the two controlling factors that aim to prevent a diversified learning path for the learner. In order to prevent too many cross-levels or too many courses in the proposed learning solution, a high value in either the $\sum\text{level}$ or $\sum\text{course}$ will be penalized through the form of a lower ranking. While more levels or courses do enrich the learning experience by presenting different applications, learning context and content presentation, it may distance the learner from the solution and may cause the learners to lose their focus when the proposed solution becomes too diversified. Also, more courses or levels are usually taken in a negative sense by the learner as they often equate more courses or levels to a longer learning curve. Hence, in their haste to master the concepts in the shortest possible time, more courses and levels are actually seen as a hindrance (instead of an enhancement) to the learning experience (learner's perspective).

4.3.2.3 Ranking Procedure

Procedure BasicQueryRank()

1. **INPUT** queryTerms, weightings, domain, constraints
2. **IF** no priority in queryTerms **THEN**
3.
$$\text{weighting}_i = \frac{1}{\text{NoOfQueryTerms}}$$
4. **ENDIF**
5.
$$\text{Vector}_{\text{query}} = \left\{ \begin{array}{l} \text{queryTerm}_1, \text{weighting}_1 \\ \dots\dots\dots \\ \text{queryTerm}_n, \text{weighting}_n \end{array} \right\}$$

6. **FOREACH** course document in domain
7. **DO**
8. sim()
9. **ENDFOR**
10. Vector₁ = {Course Document with top 50 sim() value}
11. **IF** (average of top 10 sim() value < 0.5) **OR** (Explicit multiple courses selection)
 THEN
12. AdvancedQueryRank(Vector_{query}, Vector₁, domain, constraints)
13. **ELSE OUTPUT** Vector₁

List 4.2 Procedure for Basic Search Retrieval

Procedure AdvancedQueryRank()

1. **INPUT** Vector_{query}, Vector₁, domain, constraints
2. Initialize Vector_{result} where Vector_{result} will take the format of
 Vector_{result} = {sim(), rank, Vector_{map} ID}
 where
 - rank states the numerical value of the ranking algorithm (equation 4.4)
 - Vector_{map} ID states the associated E-Map identification number where
 Vector_{map} = {ID, C_i, C_j, assoc, essentialPreReq, suppPreReq, augPostReq, ultPostReq}
3. **FOREACH** course document in Vector₁
4. **DO**
5. Assign retrieved course document as the main course
6. Retrieve all immediate course associations (associate, essential pre-requisite, supplementary pre-requisite, augment post-requisite, utilize post-requisite)
7. Retrieve the list of course's index terms that match the query terms
8. Eliminate the query terms that are found from the initial list of query terms
9. level++
10. Run comparator (courses' index terms with remaining query terms)
11. **IF** new terms found **THEN**
12. update Vector_{map}
13. **IF** all terms found **OR** level > 5 **OR** Learner Constraints Met **THEN**
14. run sim() algorithm
15. run ranking algorithm
16. **EXIT**
17. **ELSE** Assign newly found course as the main course
18. Retrieve all immediate course associations (associate, essential pre-requisite, supplementary pre-requisite, augment post-requisite, utilize post-requisite)
19. Retrieve the list of course's index terms that match the query terms
20. Eliminate the newly found query term(s) from the list of query terms
21. level++
22. Run comparator (courses' index terms with remaining query terms)

```

23.      IF new terms found THEN
24.          update Vectormap
25.          IF all terms found OR level > 3 THEN
26.              run ranking algorithm
27.          EXIT
28.      ELSE Repeat step 16 – 26
29.      ELSE END
30.      ENDIF
31.  ELSE END
32.  ENDIF
33. ENDFOR
34. Primary ranking according to sim() value
35. Secondary ranking according to ranking algorithm

```

List 4.3 Procedure for Advanced Search Retrieval **Flowchart 4.1** Basic search retrieval Procedure

Flowchart 4.2 Advanced search retrieval Procedure

4.3.2.4 EMP-Algorithm – EMPA

EMP-Algorithm - EMPA

Usage:

EMPA is used to create the E-Map. The E-Map is created in response to the learner query. It does not take the learner's learning preferences or constraints into consideration.

Input:

$\text{Vector}_{\text{result}}$ – Profile of E-Map

Output:

Possible Learning Solutions

1. **Create new concept map:** Create the number of E-Maps to populate based on the $\sum \text{trainingSolutions}$.
2. **Create new concept map nodes, connection linkage:** For each E-Map, create the number of nodes and connection linkage.
3. **Enter node name:** For each nodes in the E-Map, enter the course names
4. **Link node:** For each pair of associated nodes in the E-Map, link the nodes and label the association.
5. **Check constraints:** For each E-Map, color the nodes that do not meet the learner's query requirements (i.e. cross-domain or cross-expertise).
6. **Create hyperlink to courses:** For each E-Map, create hyperlink to link the nodes with the course offering in the system.
7. **Assign ranking:** For each E-Map, associate it with their ranking.
8. **Present Training Solution:** Present the training solution and order the E-Map according to the rank.

Algorithm 4.3 EMPA

4.3.3 P-Map

P-Map is the advance synthesis and comparison between the E-Map and the C-Map.

4.3.3.1 PMP-Algorithm – PMPA

PMP-Algorithm – PMPA

Usage:

PMPA is used to populate the P-Map. It is based on the learner's query and employed only upon successful query. The requested course(s) will be set as the targeted knowledge point and will be the sole learning target.

Input:

$Prof_{expertise}$ – Expertise Profile of User

$Prof_{preferences}$ – Learning Preference of User

C_{list} – Course(s) requested

R_{list} – List of constraints (e.g. cost per course / module, preferred time frame per course/module, etc.)

Output:

P-Map – Possible Personalized learning routes

Methods:

1. **Identify targeted knowledge point:** The sole course in the C_{list} is retrieved and assigned as the targeted knowledge point.
2. **Initialize the PVector:** The $PVector$ will be used to store the course's relationship with other courses (only associated and pre-requisite (both essential and supplementary) courses). The $PVector$ will take the form of:

$$PVector = \{C_i, C_j, assoc, essentialPreReq, suppPreReq, level, totalEssCost, totalCost, totalEssTime, totalTime\}$$

where

- C_i represents the main course id
- C_j represents the associated course id
- $assoc$ represents the association index and can take the value of $\{0,1\}$ with 1 indicating that course C_j is an association of course C_i .
- $essentialPreReq$ represents the essential pre-requisite course index and can take the value of $\{0,1\}$ with 1 indicating that course C_j is an essential pre-requisite of course C_i .
- $suppPreReq$ represents the supplementary pre-requisite course index and can take the value of $\{0,1\}$ with 1 indicating that course C_j is a supplementary pre-requisite of course C_i .
- $level$ represents the concept map's level. The targeted knowledge point will form the first level, while its immediate related courses (associated and pre-requisite (both essential and supplementary) courses) will form the next level and so on and so fro.
- $totalEssCost$ represents the summation of all the course s' fees in a particular learning route. Only courses listed as the essential pre-requisite will be taken into account. All other courses while used to aid in the synthesis of C-Map with the $PVector$, are not essential learning concepts that must be taken. Thus, the learner can choose not to take these courses and hence, these courses will not be taken into consideration until the linkage is established.

- $totalCost$ represents the summation of all the courses' fees in a particular P-Map.
 - $totalEssTime$ represents the summation of all the courses' duration in a particular learning route. Only courses listed as the essential pre-requisite will be taken into account. All other courses while used to aid in the synthesis of C-Map with the PVector, are not essential learning concepts that must be taken. Thus, the learner can choose not to take these courses and hence, these courses will not be taken into consideration until the linkage is established.
 - $totalTime$ represents the summation of all the courses' fees in a particular P-Map.
3. **Populate the EVector:** The *EVector* stores the list of the entire learner's cognitive knowledge in the particular domain. It extracts all courses $\langle baseName \rangle$ (in the particular domain) from the C-Map. The *EVector* will take the form of:
- $$EVector = \{ C_i, CName \}$$
- where
- C_i represents the acquired course id
 - $CName$ represents the course's $\langle baseName \rangle$
4. **Populate the PVector:** The *PVector* will be populated until a linkage with the C-Map is established. However, in order to prevent an infinite searching loop, searching will stop when any of these constraints are met.
- $totalEssCost > \text{learner's pre-defined cost}$
 - $totalEssTime > \text{learner's pre-defined time}$
 - $level > 5$
- The *PVector* will be populated using the breadth-first traversal algorithm. The main course will act as the root of the P-Map and course population will stem from the root.
5. **Present P-Map**

Algorithm 4.4 PMPA

4.3.3.2 Breadth-First Traversal Algorithm

Procedure BreathFirstTraversal()

1. **INPUT** *PVector*, $lCost$, $lTime$
2. initialize $level$, $totalCost$, $totalEssCost$, $totalTime$, $totalEssTime$, $found$
3. vertex $s = \text{main course}$
4. **WHILE** $found$ is not true
5. **DO**
6. **FOR** each essential pre-requisite course in $vertex$
7. **DO**
8. Add new vertex
9. Create hyperlink to course
10. Link vertex
11. Add course's duration to $totalTime$, $totalEssTime$
12. Add course's fees to $totalCost$, $totalEssCost$
13. Update P-Map
14. Get $\langle baseName \rangle$
15. Compare with *EVector*

```

16.          IF found THEN
17.              Highlight personalized path
18.              return P-Map
19.          EXIT
20.          ELSEIF totalEssTime>lTime or totalEssCost>lCost or level > 5
21.              THEN return P-Map
22.              EXIT
23.      ENDFOR
24.      FOR each supplementary pre-requisite and associated course in vertex
25.          Add new vertex
26.          Create hyperlink to course
27.          Link vertex
28.          Add course's duration to totalTime
29.          Add course's fees to totalCost
30.          Update P-Map
31.          Get <baseName>
32.          Compare with EVector
33.          IF found THEN
34.              Highlight personalized path
35.              return P-Map
36.          EXIT
37.          ELSEIF totalEssTime>lTime or totalEssCost>lCost or level > 5
38.              THEN return P-Map
39.              EXIT
40.      ENDFOR
41.  ENDWHILE

```

List 4.4 Procedure for Breadth First Traversal

4.4 Summary

It is important to note that our discussion here considers only the novel issues of learning personalization. Other issues such as learner profiles, course structure, metadata structure and methods on how to capture learner's constraints and learning preferences, while not trivial issues, are not part of this research focus. Nonetheless, we consider all these issues as implementation (NOT research) issues and we will address them in Chapter Six.

In this chapter, we focus on the formal definitions and concepts underlying the learning personalization methodology. Specifically, we show how the methodology is realized through the usage of three concept map formats.

The first map, C-Map, stores both the academic expertise and everyday experiences. Through the usage of the expertise population rules and the two C-Map population

algorithms, we have shown how we can elicit the learner's learning abilities and present them in a concept map format. The C-Map is our first step in learning personalization and paves the way for the learning content to be customized to suit an individual learner's learning abilities.

The second map, E-Map, is essentially a graphical result of the search query. However, as opposed to existing search query that uses keyword/key phrase occurrence matching to recommend learning solution that can only point to a single course or website, the E-Map is able to present a complete learning solution in a concept map format. This involves taking into account all the learning requests even if those requests span more than one learning courses, resides at different granularity levels and/or are cross-domains, cross-expertise or cross-details. Besides the requested learning concepts, the E-Map is also able to retrieve courses that are proximally related to the requested learning concepts to form a learning map of related concepts. For clarity (and educational) purposes, the relationships (associate, essential pre-requisite, supplementary pre-requisite, augment post-requisite, utilize post-requisite) between the courses will also be displayed.

The third map, P-Map, is the final desired outcome of our learning personalization methodology. Through the advanced synthesis between the C-Map and E-Map, possible personalized learning routes can be presented to the learner. Besides taking the learning abilities of the learner into consideration, other parameters such as his learning preferences and constraints will also be considered when proposing the personalized learning routes.

CHAPTER 5

GOAL DRIVEN LEARNING MODEL

This chapter describes the Goal Driven Learning Model and presents the proposed Goal Driven Learning Methodology. The proposed Goal Driven Learning Model is developed in collaboration with Dr. Shen Zhiqi from the Nanyang Technological University of Singapore and is an extension of his generic Goal Net Model [Shen, 2003, 2007]. The extended model includes a new Mathematical Model, terminologies and graphical notations. The extension is essential to enhance the existing Goal Net Model capabilities so that it can model the learning environment accurately. The main contribution of this chapter is the presentation of a Goal Driven Learning Model that besides catering to the provision of personalized learning, can also be used as a generic solution for building e-learning systems. This chapter revisits the training scenario to show how the goal model can be used to model and assist the training and learning preparation processes.

This chapter consists of seven sections. The first section provides an introduction to the agent technology and identifies existing agent interaction modeling problems. This section concludes by justifying why the goal driven learning approach is adopted. Section 5.2 looks at the nature of goal driven learning. Next, section 5.3 introduces the Goal Net Model and identifies areas that must be extended before the existing Goal Net Model can be used in the learning context. Section 5.4 presents the proposed Goal Driven Learning Methodology. Specifically, (new/extended) terminologies, mathematical model and graphical notations are presented. The justifications and applications of the proposed modeling approach are presented in section 5.5. Through a revisit of the learning scenario, the thesis shows how the proposed modeling approach is used to devise a learning solution. Section 5.6 presents the contributions and this chapter is summarized in section 5.7.

5.1 Implementation Logics

As the proposed model is implemented using agents, a brief overview of agent technology is presented.

5.1.1 Agent Technology

Recently, there has been a growing interest for e-learning systems to be actively managed using Agent Technology [Shen, 2007]. In the attempt to cater to extensive and diversified learning demands, many existing learning systems have married the power of agent technologies with the psychological principles of learning to improve educational outcomes. Agents have been increasingly employed in the field of learning to tailor to the needs of the diversity of target audiences and to create a personalized teaching and learning environment [Avouris, 2001; Garro, 2002; Baylor, 2002; McArdle, 2005]. In general, an agent is software that acts as an agent for another as in a relationship of agency. In the context of a learning environment, the agent is often employed to help the user (learner or teacher) better use, manage, and interact with a learning application such as the Learning Content Management System (LCMS) or an in-house portal system. The agent, that is programmed to act like a human, can be authorized with the autonomy to make decisions and perform certain tasks without the need for human intervention.

Agents (or multi-agent methodology) have proven their worth in many different types of applications [Jennings, 1998]. Recently, multi-agent methodology is often used in conjunction with Web-based technologies to develop and model learning environments. This is because the methodology is able to cope well with the learning environmental demands of (1) distance communication, (2) cooperation among different entities, and (3) the need to integrate different software components. Both the weak and strong notion of agent has been employed. Weak notion of agent [Wooldridge, 1995] typically characterizes an agent by the property of autonomy, reactivity, pro-activeness, and social ability while strong notion of agent [Dennett, 1987] includes the weak agency and different forms of intentionality such as beliefs, desires, intentions, goals, commitments, etc. Some examples are: Ayala [Ayala, 2003] uses agents to aid in the construction of knowledge; Boticario et al. [Boticario, 2000] use autonomous agents to support tutors;

Álvarez and Trefftz [Álvarez, 2003] use autonomous agents to act as tutors; Koper et al. [Koper, 2003] and Bergen et al. [Bergen, 2003] use agents to guide cooperation and communication between students.

Among the numerous characteristics of the agent, the weak notion of the agent is commonly exploited and applied in the field of learning. In brief: Autonomy refers to the agent's ability to control their actions and internal states without the need for direct human interaction; Reactivity refers to the agent's ability to perceive the learning environment and to respond to the environmental changes in a timely fashion; Pro-activeness refers to the agent's ability to exhibit some goal directed behavior by initiating appropriate actions to achieve its goals irrespective of the environmental changes; Social ability refers to the agent's ability to interact with other agents to unleash the power of cooperation and work towards a common goal.

5.1.2 Agent Modeling

In order to adopt agent technology and their four governing characteristics of autonomy, reactivity, pro-activeness, and social ability, a formal model for guiding and modeling the agent interaction is required. This is important, as although the agent has a high degree of autonomy and pro-activeness – they can think for themselves to decide what, how and when to achieve a learning goal and in that process – in most cases, they need to interact with other agents (either because they do not have sufficient information or capabilities to work on the problem alone or there are interdependencies between the agents).

Often, as agent systems are complex in nature, many agent modeling techniques rely on the use of object-oriented techniques to enable the complexity to be managed by decomposition and abstraction. Also, object-oriented analysis also agrees and starts from a common (agent) perspective that views the world as a set of autonomous (agents) objects that collaborate to perform some higher level function [Booch, 1994]. Hence, a large majority of the current effort on agent modeling and development employs object-oriented methodologies to model an agent as an extended object. Some examples are automata, Petri nets, Agent UML [Huget, 2004] or finite-state machines which are either message-, state-, or task- centric; that is, these systems are typically object-oriented in

nature. Such approaches model an agent's interaction through action sequences that are usually explicitly defined in terms of finite, passive messages, states, or tasks.

However, although object- and agent- oriented approaches share many similar paradigms (i.e. both adhere to the principle of information hiding, recognize the importance of interactions, etc.), there are also some significant differences. Most importantly, objects are typically passive in nature; that is, they need to send a message before they become active. This inert nature of interaction and cooperation however, limits the flexibility and adaptability of the agents. Even with the agents' ability to reason and collaborate, they are forced to follow prescribed sequences that do not take into account real-time constraints and environmental changes. Due to such limitations, these agent-based learning systems often suffer from the lack of robustness as there are limited recovery options. Furthermore, object-orientation fails to provide an adequate set of concepts and mechanisms for modeling complex systems. While object-oriented notations are good at capturing 'objects', both within and external to the system, object-oriented programming is considered too expensive in terms of computational resources. More importantly, object-oriented methodology does not give sufficient real-time support [Axelsson, 1999]; real-time support is an important factor for modeling a learning environment. Finally, the object-oriented approaches provide only minimal support for specifying and managing organizational relationships that are beyond the static inheritance hierarchies.

Hence, although we aim to adopt agent technology to implement our approach for personalized and learner-centric learning, our research has shown that existing approaches for designing the agent interaction is in need of a modeling change. Besides, as the application of the agent model is set in a learning environmental context, the nature of the learning environment as well as the learners' needs must also be taken into consideration during the modeling process. One such consideration is the Web-based nature of the learning environment. Existing modeling techniques while optimizing the agent technology on the scale of local area networks, have not been able to get such optimization to work at web-scales [Bradshaw, 1999]. Another consideration is the nature of agent interaction. The interaction between agents in a learning environment is not as complex as in other applications where the agents often have to make other agents;

undertake a particular course of action (perform some services), modify their existing course of action (delay or bring forward certain services to prevent a conflict of interest), or come to a common agreement on a course of action. So some simplifications in the existing modeling techniques should be made. Lastly, the modeling technique must also provide a clear conceptual framework whereby the learners' needs and preferences can be translated into the design and specification of agent tasks.

Hence, before agent technology can emerge as the new paradigm for learning purposes, a new modeling method that truly complements and brings out the best of the agents' abilities (i.e. ability to effect and manage change, ability to learn, communicate and collaborate, etc.) must be investigated. Here we propose a goal-driven model that models agent interactions in terms of goals. While an agent may possess different governing characteristics, its sole existence, regardless of its characteristics, is to work towards achieving its goal. Therefore, we argue that the two distinctive characteristics of goal autonomy and execution (behavior) autonomy are both required to fully characterize an agent. Hence, this necessitates a shift in the agent modeling paradigm from an object-oriented one to a goal-oriented one.

5.1.3 Why Goal Driven Learning

A learning program that uses goals to guide its understanding and appreciation of information and its context is a great improvement over one that processes everything in equal detail. Goal driven learning is far superior to data driven or object-oriented learning as the explicit elicitation of goal can lead to a systematic sequence of tasks to be mapped. Also, the difference in priority that is assigned to different tasks provides the system with added information and focus. For example, existing learning systems that are mostly completely data-driven would process everything in detail in the hope of matching something which turns out to be relevant. However, in a goal directed learning environment, learning becomes a goal-directed transformation of knowledge whereby the system can draw inferences from its goals to help it find out what it needs to know and perform in order to achieve its objectives.

Furthermore, the learning issues involved are becoming more complex and multi-faceted over time. While many e-learning systems are trying to adopt the multi-strategy learning system's method of bringing together the numerous learning algorithms, the problem of integrating the various learning algorithms is a daunting task. Besides, it is an open problem as to how to best combine the often conflicting learning mechanisms. However, the metaphor of goal-driven learning has been proven as a fruitful way of performing such integration [Cox, 1994].

Hence, we propose learning to be driven by goals. These goals are the conceptualization of the learner's needs and will be fulfilled by a set of (learning) tasks. The tasks will be determined by the system which will decide what it needs to perform in order to satisfy the learning goal(s). Sub-goals may also be identified. These sub-goals, if satisfied, will improve the system's ability to pursue the tasks that lead to the fulfillment of the goal. The decision to create sub-goals or to reason and award priorities to several active learning goals/sub-goals will be facilitated by the reasoner. Having received a learning goal, the system will first select the context as well as what it needs to learn before performing the necessary inferences. Using the learning goals as a guide, other important information such as context, environmental knowledge, and available inputs will be gathered. The system will then combine all the available knowledge and formulate a learning plan (called the Goal Net) that is appropriate for the fulfillment of each learning need.

In the context of learning, a learning goal is an explicit specification of the knowledge, skill, procedure or task that a learner wishes to master. The learning goal can be externally provided (focal goal) or internally-generated (sub-goal), domain-independent or domain-dependent, one-time or recurrent. Learning goal is a necessary component of any learning process. Given a learning need, a learning goal is conceptualized. Based on the learning goal, the system determines what parts of prior knowledge are relevant (through an analysis of the learner's C-Map), in what form the desired knowledge is to be presented (through an analysis of the learner's learning preferences), in what sequence the desired knowledge is to be presented (through the P-Map), and how the delivering of the desired knowledge can be supported (through the use of the scaffolding framework).

5.2 Nature of Goal Driven Learning

Our research models the following nature of goal driven learning: (1) active, (2) opportunistic, (3) multi-strategic, and (4) realistic.

We view learning as an active process involving the formulation of learning goal(s), the prioritization of goals (at least into primary and secondary goals and/or sub-goals), and the pursuit of learning goals using an opportunistic, multi-strategic and realistic approach. Specifically, learning will be a guided and purposeful gathering of relevant learning materials that collectively create or reinforce understanding on any subject matter. Also, the searching through the information space must be guided by a set of well-defined yet dynamic learning goals that are capable of taking environmental or learning needs changes into account.

5.2.1 Active Learning

Active learning is the study of the closed-loop phenomenon of a learner selecting actions or making queries that influence what data are being added to the training set [Cohn, 1996]. When the actions or queries are selected properly, the data requirements for the learning problem can be decreased drastically. Some NP-complete learning problems can even be simplified to become polynomial in computational time [Baum, 1991, Angluin, 1998]. In most learning systems, the learner or even the learning itself is considered as passive and learner is often treated as a passive recipient of data to be processed [Cohn, 1996]. Sometimes, the learning process has no target or goal at all and the learning system has no sense of what it is trying to learn or why it is learning. Such a passive approach ignores the fact that, in many learning situations, the learner's most powerful tool is in his ability to gather data, act on it and to reason and influence the world it is trying to understand. Active learning is hence the study of how to use this ability effectively. The identification and pursuit of what might be termed the 'learning goals' is therefore an important aspect of characterizing a learning problem [Collins, 1993; DesJardins, 1992].

5.2.2 Opportunistic Learning

Often, pieces of information pertaining to the satisfaction of a goal are not immediately available for input. Hence, the corresponding learning goal cannot be satisfied immediately and must be suspended in time pending the availability of inputs. As the learning goals and their pre-requisite inputs and tasks are indexed in memory, it is likely that the reasoner is able to find other relevant or ‘just-in-time’ information that can also lead to the fulfillment of the learning goal. In other words, the learning goals can also be satisfied opportunistically during the course of action [Hammond, 1993]. Therefore, for opportunistic learning to take place, the reasoner must possess the ability to determine when to suspend a learning goal and when to reactivate it. Furthermore, besides waiting for the ‘actual’ inputs to be available, the reasoner must also be able to remember what it needs to learn, recognize other opportunities and seize it to learn the desired knowledge.

5.2.3 Multi-Strategic Learning

Many learning systems often rely on the usage of a single learning strategy where the learning goal is defined explicitly and not by the learner himself. However, such a learning strategy does not represent the human cognition adequately. A remarkable aspect of human learners is the fact that they are able to apply a great variety of learning strategies in a flexible and goal-oriented manner and to dynamically accommodate the demands of changing learning situations [Michalski, 1995]. Hence, the use of mono-strategy suffers from a lack of flexibility and often narrows the range of application.

5.2.4 Realistic Learning

By realistic learning, we mean that we acknowledge that not all learning tasks can be performed fully. The system has limited capabilities (for example, limited learning content resources, limited computational abilities or limited knowledge about the environment). So by realistic learning, we mean that we acknowledge that there exist certain situations or instances where a (main) goal can only be partially satisfied. In such a situation, traditional goal-oriented approaches often cannot be effective. Such traditional goal-oriented approaches to building intelligent agents only consider the absolute achievement of goals (either achieved or not achieved). However, in our learning (dynamic and continuous) domain, the absolute achievement of goal may not always be feasible or possible. Hence, we have to be realistic and be able to model

situations where a goal can only be partially satisfied. Besides such modeling techniques, avenues for contingency plans in the form of secondary goals must also be provided so that the learning needs of the learners can be properly satisfied.

5.3 Goal Net Model

5.3.1 Introduction to Goal Net Methodology

The Goal Model that is proposed by Shen [Shen, 2003] is being used as the implementation platform to provide goal-driven learning. Goal Model is based on a goal-oriented modeling methodology. Through a composite state goal model known as Goal Net, a new software modeling tool is proposed. Goal Net can serve as a goal-oriented modeling tool, an agent goal model and a multi-agent identification and organization model. It assists in all phases of the life cycle in the development of agent based applications. As an agent goal model, the Goal Net enables agents to present both behavior autonomy and goal autonomy.

The goal oriented modeling method and goal autonomous agents proposed in Shen's thesis [Shen, 2003] has enabled a new modeling paradigm for agent-oriented applications and has been employed in numerous application domains such as manufacturing, military, bio-manufacturing grid and business forecasting (see for example [Shen, 2003a], [Shen, 2005], [Shen, 2006], etc.). However, as the goal oriented modeling method is designed for use in several application domains, its basic set of graphical notations and reasoning algorithm is too generic. Hence, while the model is applicable for general agent-oriented applications, a significant amount of work has to be carried out to extend it so that it can cater to the dynamic nature of the learning environment.

5.3.2 Goal Net Model

A composite state goal model is composed of five basic objects: states, transitions, arcs, branches, and tokens. A simple Goal Net example is shown in Figure 5.1.

Figure 5.1 The basic objects of the composite state goal model

States, represented by circles, are used to represent the situation that a system or a process needs to go through before it achieves its goal. There are two kinds of state objects: Atomic states, represented by unshaded circles, accommodate a single state that cannot be split anymore; Composite states, represented by shaded circles, may be split into other state objects (atomic or composite) connected via transitions.

Transitions, represented by a vertical bar, are used to represent the tasks that must be completed to change an agent from one state to another. In other words, a transition connects one state to another and specifies the relationships between the state objects it joins. Hence, each transition must have at least one input and one output state. A transition can only be fired when certain conditions are met.

Arcs, represented by arrows, are used as connectors; i.e. to connect the states to transitions and transitions to states.

Tokens, represented by black points, are used to indicate the agent's current processing position. When a token arrives at a state, it indicates a state change on that state as well as the progress of the goal being pursued.

Branches, represented by dashed arrows, are used to represent the decomposition of a composite state.

In the Goal Net, a wide range of complicated relationships among goals is supported using transitions. A transition can represent four types of temporal relationships between the states: sequence, choice, concurrency and synchronization. Sequence represents a direct sequential relationship between one input state and one output state. Concurrency has one input state but more than one output state, and all its output states can be achieved simultaneously; Choice specifies a selective connection from one input state to more than one possible output states, and only one output state can be selected and achieved by the agent. Synchronization specifies a synchronization point from different input states to a single output state, and the output state can only be achieved when all its input states are synchronized. Figure 5.2 – Figure 5.5 shows the graphical notation for the temporal relationship.

Figure 5.2 Sequence relationship **Figure 5.3** Concurrency relationship **Figure 5.4** Choice relationship **Figure 5.5** Synchronization relationship

There are three strategies of action selection mechanisms: direct, conditional and probabilistic. Direct transitions, represented by vertical bar, indicate the input states can be transited to the output states via a fixed task or a fixed sequence of actions execution. There is not any action selection mechanism involved. The conditional transitions, represented by diamonds, indicate that the task, which makes a transition fire after completion, must be selected dynamically according to runtime conditions. Rule-based reasoning will be involved for the action selection. The probabilistic transitions, represented by hexagons, indicate that probabilistic inference will be used to select actions in an uncertain environment. Figure 5.6 shows the three types of transitions.

Figure 5.6 Types of Transitions

There are two types of arcs: triangle arrows and diamond arrows. The triangle arrows mean “or” relationship between any two triangle arrows while the diamond arrows represent ‘and’ relationship between any two diamond arrows. Figure 5.7 illustrates the types of arcs with some examples.

Figure 5.7 Types of Arcs

5.3.3 Extension of Goal Net's capabilities

The original Goal Net serves as a **generic** goal-oriented modeling tool and is a multi-agent identification and organization model. It is important for the original goal-oriented modeling framework to be based on a generic set of notations and algorithms so that its framework can be applicable to other domains such as business forecasting and e-learning as detailed in the Dr Shen's thesis [Shen, 2003]. However, each application domain has its own contextual constraints, requirements and specific modeling tools and notations. Hence, this generic framework requires some refinements and extensions to its current notations and algorithms before it can be employed to effectively and completely mimic a real-life e-learning goal.

Specifically, quoting from the original Goal Model [Shen, 2003], the following areas, requiring further research and extensions, were listed below.

- a. **Further exploration of goal measurement** to enhance the goal model and to optimize the goal selection and action selection algorithms;
- b. **Exploration of various application domains** of goal-oriented modeling and goal autonomous agent/multi-agent systems, especially in personal mobility and games;
- c. **Extension of the goal model for negotiation** between goal autonomous agents.
- d. **Extension of graphical notations** for modeling e-learning-specific requirements

Working in collaboration with Dr Shen, the contributions are detailed in the following sub-section.

5.3.3.1 Clarification between State and Goal

The success and correct application of the goal oriented modeling technique hinges strongly on the explicit definition and use of the goal notation. However, in the goal oriented modeling technique, the use of 'goal' is used interchangeably with 'state' and has led to confusion. Both state and goal have the same graphical notation. The condition(s) whereby a state becomes a goal or when a goal is a state is not explicitly defined. For example, a goal is defined by Shen [Shen, 2003] as a desired state which the agent intends to achieve while no explicit definition of a state is being provided. The

closest to a definition of a state is linked to the definition of a ‘state object’ which provides a set of variables that define the profile of the state object.

5.3.3.2 Clarification on State

Continuing from the previous discussion, the model has some contradicting expectations on the use of a state. While the model claims that all jobs/functions are performed only in the task list and the state is just a ‘place holder’ for the agent without any processing capabilities, the definition of a ‘state object’ states that a state has a set of internal functions that defines behaviors on state [Shen, 2003, pp. 45].

5.3.3.3 No Mathematical Model to model Partial Goal Achievement

The original research work illustrates goal net capabilities to include goal measurement functions such as duration function, distance function and worth function. However, it is not implemented. The only form of goal measurement or achievement is simply based on selecting the goal net with the least number of goals (algorithm is based on Dijkstra’s algorithm [Dijkstra, 1959]). Also, there is no mathematical model to model partial goal achievement. Under the original goal net, a goal is either satisfied or not satisfied. However, in a learning context, the absolute achievement of a learning goal may not be feasible or even possible. There exist situations where a learning goal can only be partially satisfied. Therefore, this necessitates the need to extend the goal model to include a mathematical model to quantify the value of the ‘actual’ goal achievement. The worth value of alternative path must also be designed. This refinement contributes in the area of **goal measurement**.

5.3.3.4 Extending the Graphical Notations

The goal oriented modeling technique suffers from a lack of graphical notations. The existing notations are not sufficient to model the learning environment. Also, some notations are ambiguous in nature and need to be clarified.

- Same graphical notation used to denote different applications of state
The existing model does not differentiate the priority that is being awarded to the states graphically. Graphically, all states are being depicted by the same notation. More importantly, there is no graphical differentiation between a state and a goal.

- Insufficient information on the usage of arc

The usage of the arrows (diamond and triangle) to depict the branching of the output arcs cannot uniquely identify the exact number of output states a task list can transit to. Existing notation only represents the possible combination of output states. However, it does not provide any indications or limitations on the exact number of permissible output states.

- Additional graphical symbols are required to model the triggering of task lists

Similar to the previous argument, the current model lacks a set of comprehensive graphical notations to clearly depict the conditions required to trigger a transition from a state to a task list. Currently, the conditions to trigger a task list are denoted by triangle arrows and diamond arrows; a triangle arrow is used to denote choice transition, while a diamond arrow is used to denote direct/mandatory transition. However, no explicit procedure governs the sequence or defines what type of combination of the arrows (both triangle and diamond) is required to trigger a task list.

Figure 5.8 Ambiguity in current Goal Net Model

For example, although there is no ambiguity when only diamond arrows are being employed as shown by Case 5.1 and 5.2, the same cannot be said when triangle arrows are involved; i.e. Case 5.3 can mean (b and c) or (a and b and c) while Case 5.4 can mean c, (c and a), (c and b) or (a and b and c).

- Task List

Current task list is denoted by a single graphical notation. However, there is no indication of the capabilities of the task list. A task list can either be passive or

active. Therefore, it is necessary to use a different notation to differentiate the processing capabilities that are awarded to a particular task list.

5.3.3.5 More information required to trigger a transition firing

Negotiation between goal autonomous agents is very important in a complex, dynamic environment. Often, the satisfaction of a goal is beyond the capabilities of one agent. Hence, numerous agents have to work together to satisfy a goal. Negotiation between agents, in the original goal net, is implemented through transition firing. Four basic relationships are used; sequence, choice, concurrency and synchronization.

However, the original goal net model suffers from the lack of clarity. Using the synchronization relationship as an example, the original goal net model only states that a synchronization relationship specifies a synchronization point between states and defines concurrent relationship between states. However, it does not clearly state the requirement for goal commencement; that is, how many prior goals must be completed to trigger a commencement of a new goal. The same argument applies for the choice relationship. This refinement contributes in the area of **goal negotiation**.

5.4 Proposed Goal Driven Learning Methodology

The Goal Driven Learning Model offers a method for realizing an agent's goal through a series of adopted actions. The model is state oriented as the agent is modeled in a state oriented environment. Hence, the agent's environment consists of a finite set of states.

The Goal Driven Learning Model must address the following requirements:

1. The model must be described by its agent behavior in terms of normative relations (i.e. what the agents are obliged, permitted, or have the right to do);
2. All states that exist in the model must be based only on externally observable actions and state of affairs;
3. The model takes in three parameters: temporal situation (to identify how the situation changes and evolves over time), focal goal and its associated goal net.

5.4.1 Terminologies Defined

5.4.1.1 Goal

[Definition 5.1] A **Goal** is defined as a desired state that the agent intends to achieve. An agent will always seek to fulfill its obligation (limited only by its own processing capabilities) to reach the desired state by going through a series of state transitions from its current state to its desired state. A Goal specifies what task it wants to achieve but not how to achieve it. The explicit realization of the goal will be accomplished by the goal's associated goal net(s) via state transitions and/or agent cooperation.

Traditional goal-oriented approaches to building intelligent agents only consider the absolute achievement of a goal (either achieved or not achieved). However, in our learning (dynamic and continuous) domain, the absolute achievement of a goal may not be always feasible or possible. There exist certain situations or instances where a goal state can only be partially satisfied. The traditional goal-oriented approach cannot effectively model such a situation. Hence, we introduce an important aspect of a goal known as the completeness index. The completeness index of a goal state measures the ratio of the ideal to actual goal fulfillment in terms of percentage. Besides the completeness index, other dynamic parameters of a goal includes: an achievement measurement, a complexity index, and an uncertainty index.

- Achievement Index represents a value that captures the recognizable benefit of reaching a goal. It ranges from $[0,1]$, with 1 representing the maximum satisfaction level of reaching a particular goal;
- Completeness Index represents a percentage of the entire goal fulfillment;
- Uncertainty Index represents a percentage of the situation variables awareness;
- Complexity Index represents the required effort to realize a certain percentage of the goal. The complexity index is computed with respect to the completeness index and uncertainty index. It covers four other parameters that include (1) total time required, (2) total effort per man-hour required, (3) total cost per (sub-) goal and (4) total distance (number of states required).

These 4 dynamic parameters together with the set of (current) situation variables and user inputs will form a goal profile. This goal profile will completely characterize a Goal at any given time instance.

Hence, a **Goal** is a set of variables that effectively constructs the desired state that an agent wants to achieve and is defined as a tuple:

$$\text{Goal} = \{A_c, C_o, U_n, C_p, S_s, I_n, \text{GN}\}$$

A_c is the achievement index;

C_o is the completeness index;

U_n is the uncertainty index;

C_p is complexity index;

S_s is the set of current situation variables;

I_n is the set of user inputs;

GN is the associated goal net;

Characteristics of a goal:

1. Specific – must be defined in an authoritative way to guide the agent
2. Realistic – must be realized in a practical way by the agent
3. Desirable – must conform to the learning needs of the learner
4. Measurable – must be able to be quantifiable
5. Workable – must be associated with a plan to realize the goal
6. Temporal – can exist only at a certain time frame
7. Autonomous – must be able to act without the need for human intervention and has some kind of control over its action and internal states;
8. Challenging – must be rewarding
9. Interactive – must be able to interact with other goals
10. Completeness – must contain all the information necessary for goal fulfillment

5.4.1.2 State

[Definition 5.2] A **State** is defined as the set of proposition/state of affairs that is true at a given time t . It is one of the possible conditions an environment may exist. A State is represented by the values of the properties of the agent at a given instance. It is a unique snapshot that is quantified by the values for a set of environment variables and characterizes the agent's environmental status for that period of time. No other form of processing except for its internal processing [status check (so as to trigger transition firing) and self-assessment processing] is performed when an agent resides in a state.

The agent will constantly assess its status and will remind in its current state infinitely (within its stated permitted time frame) until it satisfies the prerequisites for state transition. Upon satisfying the prerequisites for a state transition, the agent will fire a state transition; i.e. leave its current state and progress to a processing state whereby the tasks listed in the task list will be performed. The agent will transit to the next state only upon fulfillment of the transition condition (Note that the successful execution of all the tasks listed in the task list does not necessitate a state transition).

Thus, a state can only change from its antecedent state to its directed consequent state via state transition. Through a series of state transitions, the system is able to determine whether a Goal can be realized. At any point of time, two or more states of affairs can be true; that is, a system can be in more than one states (e.g. for concurrent tasks). The profile of a state contains information about the state, including its ID, name, description, status, worth value and temporal constraints.

Hence, a **State** is defined as a tuple:

$$\text{State} = \{\text{ID}, \text{N}, \text{Des}, \text{S}_t, \text{WV}, \text{C}_t\}$$

- ID is the unique identifier for each state;
- N is the name of that state;
- Des is a short description describing the profile of the state. It includes the current state's possible antecedent and consequent states, as well as its purpose.
- S_t represents the status of the state at the current point of time. Possible statuses are focal, optional, mandatory, and compensating;
- WV is the worth value of the state at the current point in time. It is calculated dynamically based on the state's status as well as its relevance to the completion of the intended goal;
- C_t represents the set of temporal constraints that are required to fire a transition. Examples of constraint include effort per man-hour, time, and/or cost;

5.4.1.2.1 State Internal Components

While a state does not perform any (external) task execution, it has a set of internal functions that define the behavior of the state. Three sub-sets of internal functions form

the core of a state. They are the situation sensing component, the state satisfaction assessment component and the state profile component.

- Situation Sensing Component

The state must constantly assess the environment to identify when it can trigger a transition. If the set of temporal constraints are satisfied, the agent can then seize the token and fire a transition.

- State Satisfaction Assessment Component

Once an agent transits to a new state, the agent will calculate its satisfaction level of reaching the state. This value, known as the Worth Value (WV), represents the state's relevance to the completion of the intended main goal. It ranges from $\{0, 1\}$ and is calculated based on the percentage of total tasks that is expected to be completed against the total tasks that are actually completed.

Mathematical Model for Worth Value Calculation:

We propose a mathematical model for the Worth Value expression as given by

$$WV = \sum (SV_m \bullet M_i \bullet MS_i) + \sum (SV_m \bullet M_i \bullet (1 - MS_i) \bullet C_i \bullet CS_i) + \sum (SV_o \bullet O_i \bullet OS_i)$$

where:

- WV is the worth value of achieving a sub-goal. Range: $\{0 - 1\}$;
- SV_m is the satisfaction value of each mandatory state. Range: $\{0 - 1\}$;
- SV_o is the satisfaction value of each optional state. Range: $\{0 - 1\}$;
- MS_i is a boolean value (either 0 or 1) representing whether a mandatory state has been successfully achieved
- CS_i is a boolean value (either 0 or 1) representing whether a compensating state has been successfully achieved
- OS_i is a boolean value (either 0 or 1) representing whether a optional state has been successfully achieved
- M_i is the completeness index that measures how complete the tasks leading to the mandatory state have been achieved. Range: $\{0 - 1\}$;
- O_i is the completeness index that measures how complete the tasks leading to the optional state have been achieved. Range: $\{0 - 1\}$;
- C_i is the completeness index that measures how complete the tasks leading to the compensating state have been achieved. Range: $\{0 - 1\}$;

Note:

1. $\sum SV_m + \sum SV_o = 1$ where SV_m will form bulk of the percentage
2. SV_m and SV_o is the satisfaction value of fully achieving the states
3. M_i , O_i , and C_i are the runtime assessment of the percentage of tasks achieved.

This assessment will be carried out by the agent.

Assumption:

1. Each state can be associated with an explicit value that represents the satisfaction gain from realizing it.
 2. The summation of individual states worth value represents the total satisfaction gain from realizing a particular goal (focal state).
 3. Each mandatory state has its own compensating state.
- **State Profile Component**
The profile of a state contains information about the state, including its ID, name, description, status, worth value and temporal constraints

5.4.1.3 Transition

[Definition 5.3] A **State Transition** is defined as a state change function which, given its current state (S_t), and the set of current externally observable events ($Events_t$), produces a set of propositions that holds true in the next state (S_{t+1}) such that:

$$\text{State Transition: } S_t \cdot Events_t \rightarrow S_{t+1}$$

It is a response to an input task list that may cause a change of state. Transition can only occur when an agent satisfies the set of pre-conditions stated in the task list. The state of the goal can only transits from its antecedent state to its directed consequent state.

$$ST = \{Q, \Sigma, R_\Sigma, P:Q, q_o, q_f\}$$

Where:

- Q is an at most countable set of states (at least 2 states – an input and an output state);
- Σ is a finite set of elementary actions or events listed in the task list;
- $R_\Sigma = \{ \xrightarrow{\alpha} : \alpha \subseteq \Sigma \}$ is a set of relations on Q . The type of R will be refined by its own type of transition;

- $P:Q \rightarrow P(AP_i)$ is the state output map where $AP_1, AP_2, \dots, AP_n$ represents sets of atomic proposition that is true at a given instance, t . The state output map¹ takes each state to the set of atomic propositions satisfied by the state (i.e. $P(q) \subseteq AP$);
- $q_o \subseteq Q$ is the initial state (must only be one);
- $q_f \subseteq Q$ is the final state (must be at least one).

Note: In the above definition, if $\alpha \subseteq \Sigma$ and $q_o, q_f \subseteq Q$, then $q_o \xrightarrow{\alpha} q_f$ means that a state transition is successfully fired and completed and the agent has moved from state q_o (initial state) to state q_f (consequent state) by executing a list of elementary actions α . Any transition relation $\xrightarrow{\alpha}$ can be viewed as a function $f(\alpha): Q \rightarrow P(Q)$. The function $f(\alpha)$ maps q_o to the set of states reachable from q_o via a single α transition.

5.4.1.4 Task

5.4.1.4.1 Task List

Task list contains a list of tasks that should be completed in order to transit from one state to another. The successful completion of all the tasks in the task list however does not guarantee a state transition. The full set of temporal constraints must also be satisfied before the agent can seize the token and fire a transition. In special situations, the stipulated tasks need not be executed completely or successfully to warrant a transition. Several sub-task lists can also be contained in a (main) task list. Each task list will have at least one high level task. Each task will be made up of seven parameters: Name, description, priority, worth value, status, class object and resource.

5.4.1.4.2 Check List

Check list is a list of conditions that must be satisfied before a state can fire from one (either atomic or composite) state to a composite state. No tasks are performed. Check list is usually used as a predecessor to a composite state.

5.4.1.4.3 Recursive List

¹ In the temporal logic setting, the state output will be the set of atomic propositions satisfied by a state.

A recursive list is a list of tasks that must be performed repeatedly until certain conditions are satisfied. A recursive list must have a start and terminate state. The processing power associated with a recursive list is usually active in nature (see section 5.4.3.1).

5.4.2 Graphical Notations Defined

5.4.2.1 State

Figure 5.9 Actual State Representation

The state is represented by a circle that is divided in three portions to depict the three different internal components: situation sensing component, state satisfaction assessment component and the state profile component. However, for simplicity, the internal component segregation is usually not shown and a state is represented as an empty circle.

Figure 5.10 Simplified State Representation

5.4.2.2 State Granularity

5.4.2.2.1 Atomic State

Figure 5.11 Atomic State

An Atomic State is a single state that cannot be split anymore. To transit from a state (either atomic or composite) to a task list (and subsequently to an atomic state), an agent must first satisfy all the pre-requisites of its associated task list before the state can progress to the task list. The task list that is associated to an atomic state must be concrete tasks that are to be performed. These tasks must also be within the capabilities of a single agent; that is, no other form of agent interaction is necessary for the realization of the task. Atomic state is represented by an empty circle. A concrete task list is always a predecessor to an atomic state.

5.4.2.2.2 Composite State

Figure 5.12 Composite State

A Composite State encompasses 2 or more states. It can be composed by either atomic states, composite states or a combination of both. It must be realized by its associated goal net via state transitions. The goal net can be nested. The (task) scope of a composite state is outside a single agent's direct control; its realization requires a combination of atomic states that act together to perform a function or task that a single agent cannot handles. Composite state may exist in a hierarchy; i.e. a higher level of composite state can be decomposed into lower-level states connected via state transitions. A Composite State is represented by a shaded circle. The predecessor task list to a composite state is a check list that ensures that all preconditions to the realization of the composite state are met before the agent is allowed to retrieve the composite state's associated goal net.

5.4.2.3 Types of State Status²

As a State is defined as the set of proposition/state of affairs that are true at a given time t , different significance is assigned to a particular state at each instance of time. For example, when a particular state is the desired (final) situation to be achieved by an agent, that state will hold the status of a 'Goal'. Listed below are the five possible statuses of a state. Each state will hold a unique status at any given instance of time; i.e. a state cannot have more than 1 status at any given instance or situation.

5.4.2.3.1 Focal Status

Figure 5.13 Focal State

A state will be empowered the status of 'Focal' when the state becomes the goal³ that the agent intends to achieve. A 'Focal' state is the root of its associated goal net. It is represented by a 'f' to the right of the state notation.

5.4.2.3.2 Optional Status

Figure 5.14 Optional State

² A state's status is context (domain) dependent. Depending on the context and the agent's current priority, a state can be assigned different status by different agents.

³ Note that a goal is also a state. The difference between a goal and a state lies in the significance that is attached to the state at a given point of time.

A state will be assigned the status of ‘Optional’ when the achievement of the state has no significant bearing on the final outcome of the goal achievement or worth value. ‘Optional’ States are usually classified under the situations which belong to the good-to-have type. While the achievement of the ‘Optional’ state enhances the completeness of the goal, the failure to achieve it will not significantly hinder the satisfaction of the goal. However, for a more complete solution, it is recommended for the learner to plan for such states. It is represented by a ‘o’ to the right of the state.

5.4.2.3.3 Mandatory Status

Figure 5.15 Mandatory State

A state will be assigned the status of ‘Mandatory’ when the failure to achieve such state will impede the goal achievement. Such states usually lie in the main flow of the goal net where its achievement is compulsory. To ensure a higher possibility of goal achievement, it is recommended that each mandatory state be accompanied by a ‘Compensating’ state. It is represented by a ‘m’ to the right of the state notation.

5.4.2.3.4 Compensating Status

Figure 5.16 Compensating Mandatory State

Figure 5.17 Compensating Focal State

‘Compensating’ states are the alternative solution to the achieving of mandatory states. Should an agent fail to achieve the mandatory state within the stipulated temporal constraints, it will make up for it by trying to achieve its accompanying compensating state. However, these states have a worth value that is less than the mandatory states. While the achieving of compensating state is a viable (and often easier) option towards realizing the main goal, it is at the expense of a lower satisfaction level. Hence, the agent will always strive to achieve the mandatory states. ‘Compensating’ states is represented by a ‘c_m’ to the right of the state notation. One special case of compensating state occurs when a compensating state acts as an alternative solution to the focal state (goal). When this occurs, the compensating state will also be termed as a secondary goal and

represented by a 'c_f' to the right of the state notation. The secondary goal will have its own goal net. This goal will only be instantiated by the agent only when the agent fails to achieve any mandatory state as well as its associated compensating state. In such cases, the agent will instantiate the secondary goal and its associated goal net.

5.4.2.3.5 Choice Status

Figure 5.18 Choice State

A choice state occurs when an agent has the option to choose the next state to transit to. This situation occurs for conditional and probabilistic transition where any one of the 'C' states can be true depending on the runtime conditions. It is represented by a 'c' to the right of the state notation.

5.4.2.3.6 Recursive Status

Figure 5.19 Recursive-Start State

Figure 5.20 Recursive-Terminate State

R_s states the starting state whereby each cycle of recursive transition will begin from. R_t states the ending state whereby a recursive transition will end. Both states, R_s and R_t, will be branched from a task list (symbolized by) which will determine the next allowed state. Reasoning can take the form of either rule-based or probabilistic-based.

5.4.3 Processing Power of Task List

5.4.3.1 Active Task List

Figure 5.21 Active Task List

An active task list, as opposed to a passive task list, has its own processing power and hence each active task list when initialized will have its own processing thread (in a programming sense). When an active task list is initialized, a separate thread (of processing) will begin; that is, a new (task) agent will be created solely to attend to the

task list. When the (system) agent has performed the necessary actions (tasks) to trigger a transition from the task list to a new state, it will leave the task list and move on to the new state. However, as opposed to a passive task list where the task list is destroyed, the active task list can continue processing even after the (system) agent has left the task list (depending on the nature of the tasks). The processing in the task list will now be taken over by the (new) task agent. Depending on the stated tasks in the task list, the runtime environment and situation variables, the task agent will perform its own reasoning to carry out its obligation. The tasks may be performed recursively depending on the task list objective. Upon the fulfillment of its obligation, the agent will destroy its own thread and task list. An active task list will be represented by a vertical block sandwiched between 2 or more states. The block symbolizes a combination of passive tasks (vertical thin line) that is to be performed.

5.4.3.2 Passive Task List

Figure 5.22 Passive Task List

A passive task list will not hold a separate processing thread. When an agent fires a state transition from a state to a passive task list, the processing thread (agent) will move from the state to the task list. Upon fulfillment of the conditions for state transition, the task list will be 'killed' and the processing thread will move from the task list to the consequent state; that is, the task list together with all the data will be destroyed and the agent cannot return to the same task list again unless via a new state transition (as in a recursive transition). Even if the agent comes back to the same task list, the agent will treat the task list as a new task and all previous data or processing information will be lost. A passive task list will be represented by a vertical line sandwiched between 2 or more states.

5.4.4 Types of State Transition

Each transition contains of a set of pre-conditions, task list, a set of time-out alternatives (including the compensating state) and a set of post-conditions. The set of pre-conditions is the prerequisite that the agent must satisfy in order to enter the transition phase. Upon satisfying the prerequisite, the system will fire the transition and the agent will perform

the tasks listed in the task list. The set of post-conditions and time-out alternatives define the criteria for the agent to exit the transition and move on to the next state.

Note: the symbols employed in the following sub-section follows the State Transition Equation discussed in section 5.4.1.3.

5.4.4.1 Direct Transition

Figure 5.23 Direct Transition

Direct Transition indicates that the input states can be transited to the output state via a fixed task or a fixed sequence of actions execution. This type of relationship designates a direct connection in sequence from the input state to the output state. It defines successive relationship between the input and output states. The transition will fire if and only if the agent has all the necessary information and processing power to execute the tasks listed in the task list. All tasks will only be executed once and there is no action selection mechanism involved.

For Direct Transition, the following constraints will be imposed on ST.

1. $\alpha = \sum$

That is, for the transition to fire to its next state, it must complete all the tasks (functions and operations) listed in the task list. Each task will be executed completely and once.

2. Multiplicity on q_0 and q_f can only be 1

That is, there is only 1 initial state and 1 final state.

5.4.4.2 Conditional Transition

Figure 5.24 Conditional Transition

Conditional Transition indicates that the tasks, which make a transition fires, will be selected dynamically according to runtime conditions, Rule based reasoning will be used for the action selection. ‘ m ’ denotes the multiplicity index and represents the maximum numbers of allowed states that satisfies the rule constraints. For example, a multiplicity of 2 will indicates that the system is allowed to hold 2 concurrent states at any point of time.

For Conditional Transition, the following constraints will be imposed on ST.

1. $\alpha_i \subseteq \Sigma$ where $i = 1, 2, \dots, m$

That is, for each transition to fire to its next state, it must complete the entire set of task listed in its α_i .

2. α_i belongs to the set of binary relations on Q where each set of relations are rule-based.
3. m

m determines the maximum number of possible sets of α that must be satisfied at a given time t . Each completion of tasks listed in α will lead to a different final state (q_f).

4. Multiplicity on q_0 is 1
5. Multiplicity on q_f is at least 1 and the constraints on the maximum multiplicity is determined by m .

5.4.4.3 Probabilistic Transition

Figure 5.25 Probabilistic Transition

Probabilistic Transition indicates that the tasks, which make a transition fire, will be selected dynamically according to probabilistic inference. This is used in situation where a certain degree of uncertainty is present. ‘ m ’ denotes the multiplicity index and represents the numbers of allowed states that satisfies the probabilistic constraints.

For Probabilistic Transition, the following constraints will be imposed on ST.

1. $\alpha_i \subseteq \Sigma$ where $i = 1, 2, \dots, m$

That is, for each transition to fire to its next state, it must satisfy 2 conditions:

- Complete the entire task listed in its α_i .
- Meet the probabilistic assessment

Once the entire task is completed, a probabilistic assessment will be conducted. Constrained by m , any q_f that satisfies the probabilistic assessment can be a possible candidate for the next state.

2. α_i belongs to the set of binary relations on Q where each set of relations are probabilistic-based
3. m
 m determines the number of possible sets of α that must be satisfied at a given time t . Each completion of tasks listed in α will lead to a different state (q_f).
4. Multiplicity on q_0 is 1
5. Multiplicity on q_f is at least 1 and the constraint on the maximum number of permissible states is determined by m

5.4.4.4 Recursive Transition

Figure 5.26 Recursive Transition

Recursive Transition is similar to Direct Transition except that the tasks listed in the task list are to be performed recursively until certain conditions are met.

For Recursive Transition, the following constraints will be imposed on ST.

1. $\alpha = \sum$
 That is, for the transition to fire to its next state, it must complete all the tasks (functions and operations) listed in the task list. Each task will be executed completely and recursively until the exit condition is satisfied.
2. Multiplicity on q_0 and q_f can only be 1
 That is, there is only 1 initial state and 1 final state.

5.4.4.5 Jump Transition

Jump Transition, as compared to all the other transition types, does not connect the antecedent state to its directed consequent state. This is the only transition that does not require the agent to perform all the tasks listed in the task list before it is allowed to move from its antecedent state to its consequent state. Once certain criteria have been fulfilled, the agent is allowed to jump from its current state to its target state. 't' and 'c' denotes terminate and continuous respectively.

5.4.4.5.1 Jump-Terminate Transition

Figure 5.27 Jump_t Transition

Jump_t Transition indicates that once certain criteria have been satisfied, the agent will jump to its target state and the current processing will cease.

For Jump_t Transition, the following constraints will be imposed on ST.

1. $\alpha \subseteq \Sigma$ or $\alpha = \Sigma$

That is, for the transition to fire to its next state, 2 scenarios may exist. During the execution of the task list in α , the agent will constantly assess the termination condition. Upon the satisfaction of the termination condition, the processing of α will cease and the agent will jump to a pre-determined state (not directed). If the termination condition is not satisfied and the agent has finished the processing of α , then the agent will state change to its directed state. Hence, depending on the satisfaction of the termination condition, each task in α may or may not be executed completely before a state change is allowed.

2. Multiplicity on q_o is 1
3. Multiplicity on q_f is 1

5.4.4.5.2 Jump-Continuous Transition

Figure 5.28 Jump_c Transition

Jump_c Transition will continue its processing even though the agent has moved to another state. Upon completion of all the tasks, the agent will move from the current state to the directed target state (similar to Direct Transition). Hence, for Jump_c Transition, two separate processing threads (and agents) will be required.

For Jump_c Transition, the following constraints will be imposed on ST.

1. $\alpha_i \subseteq \sum$ where $i = 1, 2, \dots, m$

That is, for the transition to fire to its next state, it must complete a set of tasks (functions and operations) that is listed in the individual α_i . While the satisfaction of tasks in α_i permits a state change (non-directed), all the tasks listed in $\sum$ will still be completed (to enable a directed state change)

2. Multiplicity on q_o is 1
3. Multiplicity on q_f is at least one
4. m

The constraints on the maximum number of permissible jump states are determined by m .

5.4.5 Types of Flow

5.4.5.1 Synchronous Flow

Synchronous flow specifies a multiple transitions to task list association. That is, the tasks (listed in the task list) require a combination of (synchronous) inputs from multiple states to fire a transition. For synchronous flow, 2 types of arcs are used: diamond arrow and triangle angle. The triangle arrow is used to depict the compulsory flow while the diamond arrow is used to depict the choice flow. The total number of inputs to fire the transition will be determined by a multiplicity index and modeled as:

$$m = \sum M_i \cdot y_i + \sum C_i \cdot x_i$$

- m is the multiplicity index stating the total number of inputs required to fire a transition
- M represents the compulsory flow
- y represents the boolean factor for selecting a particular compulsory flow. As the compulsory flow must always be carried out, the y -value is always set to true and has a numerical value of 1.

- C represents the choice flow
- x represents the boolean factor for selecting a particular choice flow. Depending on the agent's reasoning and application situation, x-value can be set to either true (numerical value of 1) or false (numerical value of 0). When the x-value is set to true, the choice flow will be one of the required inputs to fire a transition.

Figure 5.29 Example of Synchronous Flow

For example, using Case 5.10, if $m = 1$ at runtime, only the compulsory input of 'c' is required to fire the transition. For $m = 2$ at runtime, a total of two inputs (either ('a' and 'c') or ('b' and 'c')) is required to fire a transition.

5.4.5.2 Asynchronous Flow

Asynchronous flow follows the syntax of synchronous relationship except that the transition need not wait for all the inputs to be present before it can be fired. Graphically, in an asynchronous relationship, the task list will not be tagged with the 'asyn' symbol. This is the most common type of relationship used.

5.4.5.3 Choice Flow

Choice Flow specifies a choice connection from one state to other states. It defines a choice relationship (conditional or probabilistic transition). This indicates that an agent have more than 1 path to choose from. The action mechanisms (probabilistic or rule-based reasoning) will be used to solve the conflict introduced by this type of relationship.

5.4.5.4 Concurrency Flow

Figure 5.30 Example of Concurrency Flow

This type of relationship specifies a concurrent occurrence between the states. It defines a concurrent relationship between the states. For example, state i can fire a transition that creates 2 concurrent state objects (state $i+1$) and (state $i+2$).

5.5 Justifications and Applications of Modeling Approach

5.5.1 Justification of Approach

Our modeling approach contributes to the classical planning systems. Most classical planning systems for learning/teaching are often not transparent as the details about learning plan generation are hidden inside software components. This makes it difficult for pedagogists to understand, and in consequence, to trust them. The proposed goal driven learning methodology contributes to this area by considering different components in the planning and execution process and modeling them explicitly into goals. Through this formal framework, each goal, which represents the planning and execution process, are not only explicit but can also be easily customized according to domain knowledge and requirements. Also, the execution of the learning goals can be easily verified and validated by the domain experts through the use of checklist for increased credibility.

The proposed goal net methodology, hence, represents a plan for the successful execution of higher-level learning tasks. Depending on the level of abstraction, the goal net can 1) correspond to a fully instantiated plan with executable steps, 2) relate higher-level goals (composite goal) to lower-level goals (atomic goal), 3) represent several alternatives out of which one is selected only at runtime when system parameters are defined, or 4) leave under-defined goals or tasks and seek advice from the user.

It is important to note that **this research does not claim any form of learning goal verification or automation methodology**. Instead, the accuracy of learning goal

elicitation and generation must be performed and validated by the SMEs. Only the content experts have the relevant domain knowledge and can decide how the individual accomplishment of sub-goals, can incrementally establish a partial ordering of primitive actions, produces a total ordering of actions and constitutes a plan for learning action sequencing.

Such a method is justifiable because as opposed to multi-agent systems, the learning goal modeling and elicitation process is less complicated as it does not requires artificial intelligent or a significant amount of learning or reasoning with uncertainty. Also, the goal modeling and elicitation follows a simple sequential decision making process that employs the explicit symbolic planning as developed in AI. Often, the teaching strategies remain unchanged. The only common change is that of the teaching strategies application sequence or timing (based on individual teaching preferences) which can be easily customized by the system.

5.5.2 Common Teaching Strategies

This section presents some common teaching strategies employed. Depending on the type of learning needs, a combination of these strategies is commonly used. Often, these teaching strategies are employed by the teachers using the fundamental principles of process-oriented learning/teaching. Following a top-down approach, the conceptualization of an e-learning service/need usually commences a full cycle of activities. Next, depending on the types of services required, the teacher will then select the appropriate teaching strategies and their accompanying tasks.

e-learning Strategies	Objectives
E-learning Service	To train employee
Learner Preparation	To prepare learner for knowledge impartation
Learning	To commerce learning
Role Identification	To identify role in a project
Employee Selection	To select employee for a specific project role
Pre-Assessment	To obtain feedbacks from the employee
Learning Path Generation	To generate learning path for the employee
Course Delivery	To deliver courses
Post-Assessment	To measure the results of learning process

Table 5.1 Common E-learning Strategies and their Objectives

While there are many teaching strategies, some teaching strategies are usually applied in most learning situations. These strategies are termed the core activities of a knowledge impartation process and are shown in Figure ---.

- Profile Analysis – Collect and analyze the profile of learner; this is essential to assist in the generating of learning path
- Path Generation – Select the appropriate learning courses for the learners according to their profile and learning requirements
- Course Delivery – Negotiate with the service provider and deliver the selected course to the learner (user customization must be performed)
- Assessment – Conduct assessment and analyze the assessment results to recommend other courses or remediate training

Figure 5.31 Core Teaching Strategies

Under current practice, each teaching strategy is fulfilled by a set of learning tasks. For example, in order to perform *Course Delivery*, the teacher will perform the following activities: 1) pre-assessment, 2) starting point negotiation, and 3) format customization. These activities can in turn be made up of smaller tasks such as those depicted in Figure -. These tasks are usually carried out in a sequential manner.

Figure 5.32 Pre-Assessment Strategy

5.5.3 Methodology Application

5.5.3.1 Methodology Step-Through

This section details how our proposed methodology is able to model the teaching strategies and activities into goals. The corporate learning scenario, depicted in chapter 2 is revisited. Here, we show how the proposed approach can be used to devise a learning solution for the common corporate learning need.

Similar to the approach taken by the teachers, all tasks that are being performed by the teachers are now being performed by the agents. The activities are still behavioral processes. Each higher level teaching strategy is being converted into goal. The activities performed under each teaching strategy will be covered by the goal net. Lower level teaching strategies that were not singular task were replaced by second-level goals. These goals are in turn accomplished by their accompanying goal net.

As per current practice, the goals act on these behavioral processes to translate them into a set of action sequences. The goals can be seen as an action selector whereby the learning goals of the learner are accomplished through an inter-woven process of selecting, guiding, sequencing and adapting the behavioral processes. In short, the initial learning wishes / hopes / desires are first transformed into intentions, then into goals,

leading eventually to action (through a sequences of behavioral processes) and hopefully, to the accomplishment of goals.

5.5.3.2 Scenario Revisit

Problem Statement

Given a new project specification (or a particular role in a project), there exists a need for the e-learning system to automatically suggest a project team, measure the human resources and competence gaps and then reduce the competence gap by creating personalized learning paths to equip the team members with the relevant project skills. Moreover, the system should be able to dynamically redefine the proposed learning paths according to feedback in order to optimize the acquisition of the needed competencies.

Requirements

Based on the above requirements, the project manager can use the goal model to cater to these requirements. The goal model will become the execution plan of the agent when the agent is instantiated. In the above scenario, the main goal is to find the project members that possess the project technical requirements. However, it is unreasonable to assume that the company is able to gather a team that consists of members who meet all the project (technical) requirements at the conception of the project. Usually, there is a need to find the best fit of personnel to the requirements (possess at least a certain amount of pre-requisite skills) and train them to be project-competence. Hence, in order to achieve the main goal, some of the sub-goals are (1) to find the best-suited personnel subjected to the project and organization constraints, (2) to acquire the employee's current technical knowledge and calculate the knowledge gap, (3) to assess the employee's competence for the project, (4) to create the employee's personalized training path and train him to be project-competent and (5) to reassess the team's ability to accomplish the project at hand.

Constraints

1. The employee must be available (no project on hand) or must be available within a certain permissible timeframe;
2. The knowledge gap of the employee must not be greater than the proposed training timeframe;

3. The cost of training each employee must not be greater than a certain stipulated cost and the total cost of training the team must not exceed the budget that is set aside for the training;
4. The training materials required to equip the employee with the relevant project skills must be available.

5.5.4 Goal Net Solution

Figure 5.33 Partial Training Goal Net Solution

Given the above scenario and constraints, part of the Goal Net Solution is shown in Figure 5.33. As this is an illustration, the full Goal Net (5 levels and 12 composite goals) is not presented. The Focal goal of the agent is “*Project Members*”. The focal goal is further decomposed into a set of sub-goals, S_1 , S_2 and S_3 , connected via state transitions.

$S_1.id = 1$, $S_1.Name = \text{“Situation Variables”}$

$S_2.id = 2$, $S_2.Name = \text{“Project Expertise”}$

$S_3.id = 3$, $S_3.Name = \text{“Tentative Project Member”}$

At the 2nd level of the Goal Net hierarchy, state S_2 is further decomposed by its associated goal net into state S_{31} , S_{32} , S_{33} , S_{34} and S_{35} where:

$S_{31}.id = 31$, $S_{31}.Name = \text{“Recommend Members”}$

$S_{32}.id = 32$, $S_{32}.Name = \text{"User Profile"}$

$S_{33}.id = 33$, $S_{33}.Name = \text{"Project Constraints"}$

$S_{34}.id = 34$, $S_{34}.Name = \text{"Personalized Path"}$

$S_{35}.id = 35$, $S_{35}.Name = \text{"Assess"}$

As an illustration of the Goal Model main transition types, Figure 5.34 which depicts part of the level 3 Goal Net is used.

Figure 5.34 S_{31} Goal Net (level 3)

$S_{311}.id = 311$, $S_{311}.Name = \text{"Suggest Member"}$

$S_{312}.id = 312$, $S_{312}.Name = \text{"Explicit Request"}$

$S_{313}.id = 313$, $S_{313}.Name = \text{"Guild"}$

$S_{314}.id = 314$, $S_{314}.Name = \text{"Similar Project/SIG"}$

$S_{315}.id = 315$, $S_{315}.Name = \text{"Available"}$

$S_{316}.id = 316$, $S_{316}.Name = \text{"Approval"}$

$S_{317}.id = 317$, $S_{317}.Name = \text{"Preferred Members"}$

$S_{318}.id = 318$, $S_{318}.Name = \text{"Available"}$

$S_{319}.id = 319$, $S_{319}.Name = \text{"Tentative Members List"}$

All transitions in Figure 5.34 are asynchronous transitions. Synchronous transitions are depicted by Task List C3, C33 and C322 in Figure 5.33. Direct transition, which indicates that the agent can go through these states sequentially to reach the goal, is shown by Task List C12 and C15. Conditional choice transition with $m = 1$ states that Task List C11 can

only fire a transition to either objects S_{311} or S_{317} . Task List C_{14} depicts an asynchronous relationship whereby any one of the antecedent states can fire a state transition. Task List C_{13} depicts an active task list which possesses its own independent processing thread. For example, when the task agent in C_{13} is initialized, a separate processing thread ($agent_1$) will be assigned to assess the learning environment and its preferred path to the desired focal goal. Next, while $agent_1$ perceives the environment, another thread ($agent_2$) will be instantiated. Based on feedbacks on the environment and its preferred path, $agent_2$ will select the next course of action to take (i.e. to suggest member either by checking for any employee's explicit request to help, or search through the list of guild members or similar project list). While $agent_1$ continues to sense for environmental changes, $agent_2$ will continue its path (from State S_{312} or S_{313} or S_{314} to S_{315} to S_{316}). In the event whereby the team has not been fully established yet (the goal of $agent_2$ has not been fulfilled), and $agent_1$ senses another possible candidate, the task agent upon successful reasoning, can instantiate another agent (similar to $agent_2$) to follow up.

5.6 Goal Driven Learning Model Contributions

Using part of the e-learning system Goal Net as an illustration, this research shows that any goal can be conceptualized at run-time to carry out the various business training needs. In our example, important learning preparation processes (set in a corporate setting) as such the learner's current and target expertise identification, knowledge gap mapping, personalized learning routes generation and assessment can be well modeled by the proposed Goal Driven Learning Model. In addition, as each state has the ability to dynamically assess the existing situation and calculate its worth value (in terms of dynamic assessment of situation changes, satisfaction gained and completeness value), the satisfaction gained from the goal achievement can be quantitatively measured and compared. Hence, this effectively enables the agent to evaluate, report its progress and intelligently choose the next sub goal autonomously for reaching the final goal.

As a comparison to existing agent-modeled e-learning systems that are usually either task-, state-, or object- oriented, our Goal Driven Learning Model holds the following advantages and contributions.

1. Purposeful Guidance

Existing agent modeling techniques rely heavily on the reasoning and learning aspects of the agent to accomplish a task. However, often, with the learning process being unbounded, much of such reasoning and learning, although in a purely logical sense, may not be useful in performing the desired goal. In fact, most of the time, learning may even cause the performance of the system to deteriorate or lead to conflicts among the agents. Hence, it is necessary to limit the proliferation of learning choices. The Goal Driven Learning Model is an improvement over existing models as the learning process is constrained and guided by goals. Through a plan formulation (Goal Net), a logical sequence of tasks is formulated that ensures that the learning that occurs is always useful.

2. Measurable

Goal measurement is an important modeling aspect to goal reasoning. While a traditional agent model can only measure a goal achievement as either reached or not reached, our Goal Driven Learning Model introduces a mathematical model to quantify the satisfaction level of achieving a goal. This is important as based on this satisfaction value, an agent can decide its next move more accurately; i.e. carry on with the existing plan or to include a compensating plan (goal) to increase the overall goal satisfaction. Furthermore, the assumed goal satisfaction value also provides the learner with additional information for selecting his personalized learning path.

3. Temporal

By introducing the time dimension and taking the dynamic environment parameters into consideration, an agent can make an informed and hence, more accurate decision during its goal pursuit at real-time.

4. Flexible

The model provides the agent with different reasoning, action selection (i.e. probabilistic or conditional) and contingency (compensating plan) strategies. The model also provides considerations about the future (i.e. assumed value of the ability to satisfy a goal). Such considerations allow the agent to be better informed so as to pursue its goal dynamically. Furthermore, as the agent receives ‘future’ information

in advance, it is more flexible as it can plan ahead. Hence, it is more robust with better recovery options.

5. Decomposable

The goals can be decomposed into smaller and simpler sub-goals. Hence, the overall complexity of achieving the goal is reduced. A lower complexity will better the chance for the goal completion.

5.7 Summary

In this chapter, a practical, formal goal driven learning modeling approach for eliciting, modeling and analyzing e-learning goals is presented. In contrast to other agent-oriented e-learning systems, a goal-oriented agent based e-learning system can also be used in the area for planning. A comprehensive approach to specifying and validating the plans of the e-learning agents through the notion of goals is proposed. The goal model set the formal framework for the planning and customization of e-learning systems. The extended goal relationships and their representations are elaborated on. Furthermore, the goal measurement and properties of the model are discussed. An illustration of the use of the proposed formal goal model set in a training environment is also discussed.

CHAPTER 6

FRAMEWORK AND PROTOTYPE DESIGN

This chapter presents the documentation for the entire design process of the learning environment and prototype system. Although this research only addresses problems in the e-learning domain, the system is modeled with an architecture that is flexible enough to be applicable for many other applications.

Section 6.1 describes the complete architectural considerations. It lays a strong technical foundation to ensure that the learning environment is built upon an interoperable, robust, and flexible infrastructure that is able to support other advanced services and elaboration that will come into existence as the system matures. Section 6.2 looks at the framework design and standards that are adopted. It also explains how these standards and other technology advances can be employed in the context of learning. A harness framework of e-learning is proposed. Next, implementation details pertaining to the prototype design are presented in section 6.3. Specifically, a learning platform, termed E-Grid, is presented. E-Grid substantiate our claims and show how our proposed framework can harness the capabilities of three emerging technologies, namely, Grid Computing, Semantic Web and Web Services for e-learning. One of the implemented modules realizes the ideas and solutions presented in the early chapters of the thesis. This chapter is summarized in section 6.4.

6.1 Architectural Considerations

To cater for e-learning, the system must be open in its architecture and support a large and extensible class of distributed digital information services. The basic entities to be stored, retrieved and managed will be information stored in the form of learning objects. In the system's context, these learning objects will be seen technically as a form of digital objects to be stored, accessed and managed. Therefore, the basic function is to provide an access protocol service for locating and disseminating these digital objects.

This basic function constitutes a minimal set of requirements for the system to be fully functional. While it is not the intention to build a universal, wide-area digital information infrastructure, considerable design efforts are employed to ensure that the infrastructure, at its very least, must be interoperable, robust, and flexible enough to support other advanced services and elaborations that will come into existence as the system matures. Hence, the system must be built upon an interoperable layer that will not restrain the higher level user and service level choices such that it will become inappropriate to add in new features at any point of time.

With the primary aim set at creating a learning platform for supporting social, intellectual, and technological innovation for learning, the system will harness the capabilities of Grid Computing, Semantic Web and Web Services. These three technologies are seen as the future IT technologies that will spearhead the knowledge revolution. Currently, many current e-learning platforms and systems that have been developed and commercialized are based on client-server or on peer-to-peer [Pankratius, 2003]. More recently, Web Services-based architecture [Lin, 2006] has also been employed. However, a major drawback of such implementation platforms lies in the limitation in scalability, availability, and distribution of computing power and storage capabilities. Thus e-learning is often being deployed in areas where high requirements in any of these aspects are not mission-critical, which excludes its exploitation, for example, in most nature sciences or medical areas [Pankratius, 2003]. Therefore, Grid Computing is used to extend the environment's generality and efficiency in the sharing of resources. The design of a collaborative workspace for supporting diversified virtual, Grid computing and e-learning applications is also one of the main architectural considerations.

6.2 Framework Design

The knowledge environment will be developed using the layered architecture shown in Figure 6.1. The layered architecture is chosen over other architectural patterns such as client/server, three-tier or peer-to-peer pattern as a layered architecture aids in organizing the system into layers of subsystems. This is extremely important as organizing the design model in layers aids reuse, one of the most important aspects of e-learning. Besides reusability, portability is also enhanced, as changes are isolated only to one layer. Each layer in the layered architecture is a set of subsystems that share the same degree of generality and interface volatility.

Figure 6.1 Layered Architecture

6.2.1 Functional Layers

The knowledge framework exploits the capabilities of open-standard Web technologies such as XML and Web Services to construct a Grid infrastructure that harnesses the Semantic Web advantages. The Knowledge Framework encompasses three functional layers, namely, the wire layer, the description layer and the discovery layer.

Discovery	WSIL	UDDI	OWL+OWL Schema/DAML+OIL
Description	WSDL	GSDL	RDF + RDF Schema
Wire Protocol	SOAP / XML Protocol		
	XML		

Figure 6.2 Functional Layers

6.2.1.1 Wire Protocol Layer

The Wire Protocol layer forms the platform and specifies the wire protocol for supporting the client and server running on different machines in the distributed network. Among the numerous wire protocols, Simple Object Application Protocol (SOAP) is chosen as it is

the technology that will be deeply embedded in the future of distributed computing. Besides, SOAP offers ease of interoperability when used with Web Services.

Furthermore, with the ease of coupling with other technologies such as the Universal Discovery, Description, and Integration (UDDI) and Web Services Description Language (WSDL), SOAP is selected as the wire protocol to transform the knowledge applications and used to aid communication over the Web with the concept of Web Services. However, as SOAP is merely a wire protocol, it does not implement security. Thus, we are proposing the use of SOAP with Extensible Markup Language (XML) protocol to allow the employment of application-level security.

As SOAP is a text-based protocol that uses XML for data encoding, XML will be employed at the base of the framework. XML is a cross-platform, extensible, text-based standard for representing data. It is also a key technology for Web Services development.

The standards employed are XML Version 1.3, and SOAP Version 1.2.

6.2.1.2 Description Layer

The Description Layer lies directly above the Wire Protocol Layer. With a sound and solid platform exemplified by XML and SOAP or XML Protocol, this layer serves as the general purpose description language layer for identifying the required information. With the proliferation of information, it is extremely difficult to find the exact information required in this giant information space. This problem is further amplified when dealing with different machines in distributed networks. Besides the problem of locating and understanding the information, the proliferation of communication protocols and message formats pose another challenge. Thus, this layer is included to address such problems.

To enable efficient retrieval, the environment needs to understand the information's semantics. This is done through the use of metadata. Metadata has the potential of transforming information from 'machine readable' to 'machine understandable'. We employ Resource Description Framework (RDF) and Resource Description Framework Schema (RDFS) to achieve "machine understandable" information. Although RDF is used as the 'descriptive language' for resources, it is a further W3C initiative, building upon earlier developments such as Dublin Core and the Platform for Internet Content Selectivity (PICS) content rating initiative. Thus, while RDF is used to empower creation,

exchange and the use of metadata, the framework will employ the Dublin Core Element Set (15 elements) to capture a representation of the essential aspects related to the description of learning resources.

This layer also includes XML Namespace (XML-NS), a construct which is adopted in RDF to provide a more comprehensive system to describe all aspects of the resources. XML-NS provides a simple method for qualifying elements and attribute names by associating them with namespaces identified by URI references. It harnesses the great power of XML and RDF to declare their own modes of expression for resources description and yet allows 'authorship' sharing. In summary, the framework will employ the Dublin Core Element Set whenever possible and use XML/RDF to allow terms in a metadata vocabulary to be treated as reliable, unambiguous, describable objects.

Besides providing description, this layer also looks into describing the communication of protocols and message formats. To provide a standardized way to structure communication, WSDL is used. WSDL defines an XML grammar for describing network services as a collection of communication endpoints capable of exchanging messages. It is extensible and allows description of endpoints and their messages regardless of the message format or network. In short, for Web Services, although SOAP offers the basic communication, it does not state how the messages must be exchanged to successfully interact with a service. Hence, WSDL is used as a specification language to describe all the available services through a set of endpoints operating on the message.

Aiming to incorporate the dynamic service nature of the Grid, the Open Grid Services Architecture (OGSA) is included to bring about a convergence of the Grid and Web Services communities. As OGSA adopts the general Web Services approach and protocols to build its own Grid infrastructure, we will use Grid Service Description Language (GSDL) on top of WSDL. Grid services, described in terms of GSDL, will serve as an additional building block to function as an enhanced Web Service that extends the conventional Web Service functionalities into the Grid domain. These Grid services will be published and registered in the Discovery Layer through a specialized Grid registry and can be invoked by using Internet protocols such as SOAP.

The standards employed are RDF Version 1.0, RDF Schema Specification 1.0, DCMES 1.1 and WSDL Version 2.0.

6.2.1.3 Discovery Layer

The first two layers discussed so far are required for interoperable Grid and Web Services. They serve as the basic infrastructure to enable specific services and applications to leverage the current Grid and Web Services over the Internet. The next layer, the Discovery Layer will deal primarily with the provision, publication and discovery of existing services for knowledge discovery. These services can be classified into two domains, namely, Grid and e-learning domain, and uses the Web and Semantic Services for publication and discovery. This layer covers the interaction between machines that agree on the protocols and services described by the lower layers to aggregate the supporting and more specific services or applications. Specifically, Grid provides the ability to perform higher throughput computing by taking advantage of the many networked computers to model a virtual computer architecture that is able to distribute process execution across the learning environment. Services, in the context of Grid, are computing machines while services in the context of e-learning are learning resources.

The Universal Description, Discovery and Integration (UDDI) and the Web Services Inspection Language (WSIL) will be used as the mechanism to discover the service. Service Discovery is defined as any action that enables a service requester to find and access the WSDL documents. This process can be as simple as accessing a file or URL containing the WSDL or as complex as querying a UDDI registry and using the WSDL file(s) to select potential services.

UDDI is a specification for a registry that can be used by a service provider to publish WSDL documents. Clients search the registry looking for services and then fetch the WSDL documents. Building upon the basic platform of XML, XML-S and SOAP, UDDI provides a foundational infrastructure for a Web Services based environment for the Grid and e-learning domain.

Alternatively, the service descriptions can be published locally within the service hosting the environment and discovered using distributed protocols such as WSIL. WSIL is also an XML-based document format that allows the inspection, discovery and aggregation of

Web Services description in a simple and extensible fashion. While similar in scope to UDDI, WSIL is a complementary but different model to service discovery. Unlike UDDI, WSIL approaches service discovery in a more decentralized fashion with the assumptions that the service requester is already familiar with the service provider.

This layer also integrates the Semantic Web vision of machine interpretability of content by incorporating semantic meanings. Web Ontology Language (OWL) (<http://www.w3.org/TR/2003/WD-owl-features-20030331/>) is intended to be used explicitly to represent the meaning of terms in vocabularies and their relationships. It is built on top and beyond the basic semantics of RDF-S and will be used to formally describe the terminology. Thus, OWL goes beyond XML, RDF, and RDF-S in its ability to represent machine interpretable content.

The standards employed are WS-Inspection 1.0, UDDI, OWL, DAML + OIL.

6.2.2 Proposed Harnessed Framework

The proposed harnessed framework represents the full-fledged learning environment and other technology advances. However, as this thesis concentrates on e-learning, only the key e-learning modules, namely, the Ontology Services, Inference Services, Collaborative Authoring Services, Augment Services and Agent Services, will be discussed.

Figure 6.3 Proposed Harnessed Framework

6.2.2.1 Ontology Services

To serve as an interoperating medium for communication, explicit specification of learning resources must be established. A common understanding is vital for different

machines to use and share knowledge. This is established in the form of ontology. Ontology is an explicit specification of some topic. It is a formal and declarative representation which includes the vocabulary for referring to the terms used in a particular subject area and the logical statements that describe what the terms are, how they can or cannot be related. In our learning context, ontology is defined as follows:

[Definition 6.1] Ontology is defined as a formal explicit description and representation of the learning resources and their relationships in a domain of discourse.

It will contain vocabulary that stores machine-interpretable definitions of the concepts as well as the relations among them. Employing ontology for communicating knowledge will put the system in line with the Semantic Web Vision whereby information is given explicit meaning, making it easier for machines to automatically process and integrate information available on the Web.

With OWL employed as the ontology language, the system will use the open source Protégé-2000 tool as the system's ontology and knowledge-base editor. Protégé is an open-source, Java tool that provides an extensible architecture for the creation of customized knowledge-based applications. Although it is currently only being used in clinical medicine and biomedical sciences, it can be used in any field where concepts can be modeled as a class hierarchy. As an ontology editing tool, the system uses the Protégé's OWL Plug-in to describe the ontology declaratively and to state explicitly what the class hierarchy is and to which class it belongs.

Employing a concept-based approach, a concept will be described by a class. Classes will be the main focus of the ontology. Thus, ontology together with individual instances of classes constitutes the knowledge base. However, as ontology development and design is not the main research focus, this research will not focus on its development. Instead, we will adopt standardized libraries of Web ontologies such as the Ontolingua ontology library, the DAML ontology library and the OpenCyc Selected Vocabulary and Upper Ontology. Although there exist a number of publicly available commercial ontologies (e.g. UNSPSC, RosettaNet) and educational ontologies (e.g. DMOZ, KSL), the system also acknowledges that sometimes, no relevant ontology for a particular subject exists. Thus, the SMEs can also develop a new ontology from scratch using Protégé. Although,

there is no single correct ontology-design methodology, a rather comprehensive guide exists in Ontology Development 101 [Noy, 2001] and will be adopted.

6.2.2.2 Inference Services

The conceptual logic of the system will be implemented using the fundamental theory underlying Language Engineering (LE). LE is the discipline or act of engineering software systems that perform tasks involving the processing of human language. Both the construction process and its outputs are measurable and predictable [Cunningham, 1999]. It is the application of Natural Language Processing (NLP) to the construction of computer systems that processes language for some tasks usually other than the modeling language itself. LE will be used to create knowledge on the language (partially through the use of ontology) to enhance the query and retrieval capabilities of the system. It will enable the system to access the learning resources efficiently and focus precisely on the exact information that the learner needs, thereby, saving time and avoiding information overload. This will also undoubtedly enhance the learning experience.

The inference engine creates a conceptual domain of the query and provides possible reasoning over the conceptual representation. It then performs dynamic semantic analysis on the search criteria and matches them against the predefined semantics of the learning resources, thereby, giving a far better result than simple keyword searches.

The knowledge representation and reasoning services will be carried out using Description Logic (DL). DL has been used in a range of applications for reasoning with database schemas and queries [Calvanese, 1998]. They are also used widely as a formal basis for ontology language and to provide reasoning support for ontology design and deployment. To exploit the semantics of ontology based relations, the DL-based inference engine must be able to reason with the ontology language, OWL. Hence, the inference engine will perform reasoning with OWL through the use of OWL DL.

Besides DL, Latent Semantic Indexing (LSI) can also be used to aid reasoning. LSI [Deerwester, 1990] is a well-defined mathematical method, which uses mathematics to define the semantic cohesion between documents and collections of documents. Besides pure keyword matching, LSI also employs indexing to find semantically close concepts that are associated with keywords. For example, when a learner wishes to learn

“Quantum Mechanics”, instead of searching solely for the keyword of “Quantum”, “Mechanics” and “Quantum Mechanics”, the system will also associate semantically close concepts such as “Wavelength-Momentum Relation”, “Schrodinger’s Equation” or “Wave Functions” to Quantum Mechanics although these concepts might not contain the keyword at all.

6.2.2.3 Agent Services

The Goal Driven Learning Model that is proposed in Chapter five is implemented through this agent service. From the logic view, the agent service can be formally specified as a 8-tuple, $\{S, A, GN, K, S_s, R, FG, FA\}$, where

Agent Services = $\{S, A, GN, K, S_s, R, FG, FA\}$

S is a set of states defining agent goals;

A is a set of actions defining agent behavior;

GN is an agent goal model, i.e. Goal Net;

K is the knowledge of the agent;

S_s is a set of current situation variables;

R is a set of situation-action rules defining goal autonomy of the agent;

FG is the function for goal selection defining goal autonomy;

FA is the function for action selection defining behavior autonomy;

The agent services will follow the Goal Net Methodology and consist of the following tasks that will be repeatedly performed to form a PR_2A cycle.

- Perceive – The agent perceives the environment continuously to sense any new situation
- Reason for goal selection – The agent infers the next goal based on its goal model, its knowledge, and its perception of the environment
- Reason for action selection – The agent infers its actions based on its goal model, its knowledge, and its perception of the environment
- Act – The agent executes the selected actions

6.2.2.4 Collaborative Authoring Services

Figure 6.4 ADDIE Model

This service is designed with the basic concepts of the ADDIE Model and modified to cater to the learner-centric environment in the e-learning context. This service implements the pedagogical considerations that have been discussed in chapter four. The pedagogical considerations have been transferred and implemented through a content development model.

Figure 6.5 Content Authoring, Assembly, Management and Delivery Process

6.2.2.5 Augment Services

Knowledge augmentation will be fulfilled through the use of resource annotation. Initially, the system will employ Amaya as the Web editor to create rich annotation content for each concept – Amaya is an open source software project hosted by W3C (www.w3.org/Amaya). It is a tool that can be used to create and update documents.

6.3 Prototype Design

The prototype system is implemented with the objective of presenting the research contributions and demonstrating how the proposed theories and methodologies can be realistically implemented. As a proof of concept, the proposed learning platform that harnessed the capabilities of the three emerging technologies is implemented successfully

as a Master of Science Project in the Nanyang Technological University of Singapore. The harness framework of e-learning is termed e-Grid (e-learning Grid Network) and is detailed in Section 6.3.1. The proposed theories and methodologies that were discussed in Chapter three and four were implemented using a local host and server and is documented in Section 6.3.2.

6.3.1 E-Grid

6.3.1.1 Motivation and Proof of Concept – Harnessed Framework

This section presents an e-learning grid network called E-GRID that creates a learning community through an online network of servers and computers. The motivation for its creation is twofold:

1. To implement the proposed harnessed framework of e-learning that is proposed in the preceding section
2. To substantiate our claims that the proposed framework can harness the capabilities of three emerging technologies, namely, Grid Computing, Semantic Web and Web Services for e-learning

6.3.1.2 E-Grid Overview

E-Grid essentially leverages on existing hardware and software for sharing heterogeneous resources (based on different platforms, hardware/software architectures, and computer languages), located in different places, and belonging to different administrative domains, over a network, using open standards. In short, it involves the virtualization of computing resources and unifies different resources such CPUs and storage devices remotely. It manages the distribution of computations, storage and movement of data across computers in a grid. This is done in a transparent way so that users of the grid are not aware of it. Through this, available resources can be used more efficiently and the aggregation of idle CPU cycles can match supercomputing capabilities. To enable scalability and adaptability, Web technologies and standards such as the Extensible Markup Language (XML) and Web services are used.

6.3.1.3 E-Grid Architecture

In the e-learning platform, the grid is implemented as a middleware to provide all grid-related services. It uses the Internet as a communication infrastructure. The generalized, layered, protocol architecture for grids can be represented as shown in Figure 6.6.

Figure 6.6 Grid Middleware

The infrastructure of the EGRID system is shown in Figure 6.7. It illustrates the grid and its connection to the satellite servers and client access. For illustration purposes, only eight of the n-number of nodes in the grid is shown as satellite servers. In practice, n, the total number of computers joined in the grid as nodes, can run into thousands and even millions. The E-Grid network is successfully implemented and tested in Nanyang Technological University (NTU). The acronym of EEE, HSS, MAE and BS represents the different departments in NTU.

Figure 6.7 E-Grid Infrastructure

The EGRID system is made up of a network of servers and clients connected on the world-wide web. The primary communication protocol is the universal internet protocol, TCP/IP. The design makes use of standard existing hardware and software so that new computing resources can join the grid easily to reduce cost and complexity. Maintenance of the grid is also simplified through remote procedural calls and other standard browser-based client software.

Figure 6.8 shows one of the pages of the learning management system running within the grid. The search engine and the server re-direction are displayed showing the operation of some of the middleware features.

Figure 6.8 Learning Management System Information Page

The implemented e-learning grid architecture is made of four layers as shown in Figure 6.9. Figure 6.10 shows the comparison between the proposed model (figure 6.3) and the implemented prototype (figure 6.9).

Figure 6.9 E-Grid Architecture (Implemented)

Figure 6.10 Implemented Architecture Vs Proposed Architecture

6.3.1.4 E-Grid Modules

6.3.1.4.1 Infrastructure Layer

This is the bottom layer of the grid which is made up of all the basic network support environment devices and systems such as computing devices (servers and clients) and network protocols (TCP/IP). The use of the TCP/IP simplifies the connection process so that devices can join the grid for unlimited expansion and growth of the grid network.

6.3.1.4.2 Web Services Layer

The web services layer is for implementing the basic web services related protocols. This includes XML, SOAP, UDDI and WSDL. This layer provides the elementary connectivity, interoperability, reliability and flexibility for the layers on top of it. SOAP was chosen mainly due to its use of XML syntax, providing advantages in terms of internationalized information (via Unicode), existing and upcoming tool support (both open source and proprietary).

6.3.1.4.3 Grid Middleware

This layer is a crucial layer to build the grid environment. The main functions implemented in EGRID are:

- File Distribution
- Redirection

- Database Synchronization
- Load Balancing
- Searching Engine
- Security Services

The middleware provides the dynamic load balancing service, redirection service, file distribution service, database synchronization and security services. Dynamic load balancing (DLL) is implemented through the Apache-Tomcat active service page application server. DLL provides redundancies in the network to ensure a hundred percent availability. It also provides the persistency and server monitoring service. The re-direction service helps to reduce the main servers load by re-directing services to the nearest active server in the grid or the school server of the user. This increases the EGRID performance by distributing the bandwidth across the active nodes in the grid.

EGRID takes advantage of the node servers in the grid to increase its performance by distributing the learning content files across well defined node servers. This allows the system to render the information to the user through the nearest node. This is possible because all users are registered with the system through the center, school, course and class chain. Files uploaded by the administrator are distributed to all satellite servers while files uploaded by instructors are distributed to the respective cluster servers. As a result the load of the main servers will be reduced and the utilization rates of the other machines are increased.

The search facilities as well as the personalized the file content search engine that enables users to search for information in the grid network will be explained in detail in the next section.

6.3.1.4.4 Content Repository

The content repository is implemented according to our proposed theorems and methodologies and is explained in detail in section 6.3.2.

6.3.1.5 Implementation Packages

Figure 6.11 Implemented Packages for E-Grid

6.3.1.5.1 General Retrieval Package

This package allows the user to view the general information such as login information, logout information, course information, messages, user profiles (display portal) and to perform basic tasks such as to update personal particulars, to plan learning tasks, and participate in discussion and communication forum. This includes all user perspectives; student, instructor and administrator. The service in this package requires the user to register beforehand.

6.3.1.5.2 Customized Service Package

This package allows the user to register to the system. After successful registration, the registered user can view their profile and other general retrievals can be performed. They can also modify their profile and the system will maintain the updating of the database.

6.3.1.5.3 Basic Analysis Package

This package provides users the following functionalities:

- View File
- Upload New File
- Delete File
- View Course Details
- Add Course Details

6.3.1.5.4 Advanced Analysis Package

This package provides users with more advance functionalities and is traced out below:

- Redirect Management
- Planner Management
- User Profile Management
- Classroom Profile Management
- Discussion Board Management
- Communication Management

6.3.1.5.5 Database Administration Package

This package provides the maintenance of system database to ensure the validity of data that is provided to the user. This package traces to the following

- Replication
- Master and Slave

6.3.2 Content Management System

The content management system is implemented at the content repository level in the E-Grid system. However, this system is discussed in isolation as this is an important section as it implements the algorithms and methodologies that were proposed in the early chapters. As the network considerations such as network publishing are already realized by E-Grid, the implementation of the content management system is simplified and developed in isolation using a local host and server.

Similar to E-Grid, Java is used to implement the content management system. The system comprises of three main components and four modules.

1. MySQL

The MySQL Database, a popular open source database, is used because of its consistent fast performance, high reliability and ease of use.

2. Apache Tomcat

The Apache Tomcat is the servlet container that is used in the official Reference Implementation for the Java Servlet and Java Server Pages (JSPs) technologies.

3. E-Learning Portal

The portal is developed using (1) Servlets to encompass the business logic, (2) JSP to display the graphical interface, (3) JavaBeans to store the persistent data.

Figure 6.12 Overview of Content Management System **Fig. 6.13a** Object Authoring Module**Fig. 6.13b** Semantic Search Module **Fig. 6.13c** Dynamic Concept Map Module **Fig. 6.13d** Learner Profile Management Module**Figure 6.13** Internal Modules of Content Management System

The relationship between these components is illustrated in Figure 6.12. JSP/Servlets run within a web server called a servlet engine where Tomcat is used. They dynamically generate HTML that is streamed back to a Web browser when a JSP page is accessed. One main advantage of using these components is that the learner never has to know where the content is stored as the content is generated on-the-fly by Java. Hence, this allows a Web site to access data in a database and build complex applications. The system is implemented using four modules: Object Authoring Module, Semantic Search Module, Dynamic Concept Map Module and Learner Profile Management Module.

6.3.2.1 Object Authoring Module (OAM)

The Object Authoring Module implements the pedagogical considerations and proposed strategies that are discussed in chapter three. Key pedagogical strategies such as instructional strategies (micro-level), learning event classification (macro-level) and scaffolding framework as well as additional key personalization considerations such as learner's learning preferences, determination of the nature of knowledge to be delivered, the type of learning outcomes and learning/performance objectives sought and the identification of the pre-requisite knowledge required are realized by the OAM.

6.3.2.1.1 Learning Resource Structure

The size and scope of the learning resources will follow the five levels content hierarchy (Reusable Information Object (RIO) → Reusable Learning Object (RLO) → Topic → Module → Course) as proposed in Chapter 3, Section 3.2. Besides the content hierarchy, it is also useful to distinguish between the levels of content complexity, content details, and the type of presentation formats so that the system can adapt the learning resources according to the learner's needs. The relationship with other learning resources must also be specified to enable concept mapping.

Expertise Levels	- Novice, Intermediate, Expert
Detail Levels	- Overview, Normal, Detailed
Presentation Formats	- ILS (Active/Reflective, Sensing/Intuitive, Visual/Verbal, Sequential/Global)
Relationships	- Associate, Essential Pre-requisite, Supplementary Pre-requisite, Augment Post-requisite, Utilize Post-requisite

These parameters of the learning resources will be discussed in Section 6.3.2.4.

6.3.2.1.2 Conceptual Data Modeling

The database tables and fields are formed prior to the portal set up. Entities and attributes are identified. Entities are typed in CAPITAL letters. Attributes are typed in *italics*. As a prototype, only the learner and the learning resources at the course level is modeled.

1. A number of COURSEs are offered in the e-Learning portal. These COURSEs are classified into different *CourseLevel* (course level) namely course, module, topic, RLO and RIO.
2. A COURSE must be assigned to a minimum of three and maximum of seven internal COURSEs which are one level lower than its level. E.g. if the COURSE is created at the course level, the internal COURSEs must be at the module level.
3. A COURSE can have zero to many course associations (associate, essential pre-requisite, supplementary pre-requisite, augment post-requisite, utilize post-requisite).
4. A COURSE must be assigned to at least one *Concept* with total *Weightage* of 1.0.
5. A LEARNER registers and logs in to the portal.
6. A LEARNER can search for a COURSE based on the *Domain*, *AcadLevel* (academic level), *Concept* and *Weightage*. For advanced search, the LEARNER can cross search in other *Domain*, *Expertise* and the *SearchLevel* (search level).

LEARNERINFO

learnerId	firstName	lastName	dob	nationality	address	contactNo	email
-----------	-----------	----------	-----	-------------	---------	-----------	-------

LEARNERPIN

learnerId	userName	password	registrationDate
-----------	----------	----------	------------------

LEARNER QUESTIONNAIRE

learnerId	answers	dateTaken	ils	act_ref	sen_int	vis_vrb	seq_glo
-----------	---------	-----------	-----	---------	---------	---------	---------

COURSEINFO

cId	cName	cAuthor	cDateCreated	cCost	cLocation	cDomain	cLearningPref	cExpertise	cDuration	cDetail
-----	-------	---------	--------------	-------	-----------	---------	---------------	------------	-----------	---------

COURSE ASSOCIATED

id	courseId	aCourseId
----	----------	-----------

COURSE AUGMENT

id	courseId	aCourseId
----	----------	-----------

COURSE ESSENTIAL

id	courseId	aCourseId
----	----------	-----------

COURSE GOODTOKNOW

id	courseId	aCourseId
----	----------	-----------

COURSE INTERNAL

id	courseId	aCourseId
----	----------	-----------

COURSE UTILIZE

id	courseId	aCourseId
----	----------	-----------

Table 6.1 Relational Database

Name	Type	Records	Size	Last Update
acadlevel	MyISAM	4	3 KB	9/20/2006 6:06:37 PM
course_associated	MyISAM	139	9 KB	9/21/2006 11:58:21 PM
course_augment	MyISAM	54	4 KB	9/21/2006 11:58:22 PM
course_essential	MyISAM	61	4 KB	9/21/2006 11:58:21 PM
course_goodtoknow	MyISAM	44	4 KB	9/21/2006 11:58:22 PM
course_internal	MyISAM	294	15 KB	9/21/2006 11:58:21 PM
course_utilize	MyISAM	46	4 KB	9/21/2006 11:58:22 PM
courseinfo	MyISAM	250	15 KB	9/21/2006 11:58:18 PM
courselevel	MyISAM	5	3 KB	9/21/2006 10:43:49 PM
detail	MyISAM	3	3 KB	9/21/2006 9:45:35 PM
domain	MyISAM	84	5 KB	9/20/2006 6:06:37 PM
expertise	MyISAM	3	3 KB	9/21/2006 9:43:13 PM
learner_questionnaire	MyISAM	3	3 KB	9/20/2006 6:06:37 PM
learnerinfo	MyISAM	8	3 KB	9/20/2006 6:06:37 PM
learnerpin	MyISAM	8	3 KB	9/20/2006 6:06:37 PM
learningconcept	MyISAM	107	7 KB	9/21/2006 11:58:22 PM
questionnaire	MyISAM	44	8 KB	9/20/2006 6:06:37 PM

Figure 6.14 Screen Capture of Database

6.3.2.1.3 Course Description Language (CDL)

The metadata of the learning resources follow the Dublin Core Metadata Element Set, Version 1.1 and is meta-tagged using the 15 elements namely, Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage and Rights. A separate Course Description Language (CDL) is also tagged to the learning resources to enrich the metadata and to allow conceptual mapping. **Through CDL, the pedagogical and personalized considerations can be achieved.**

The root element of CDL is the <course> element. The usage of the term, course, gives no indication of the hierarchical level of the learning resource. The <information>, <structure> and <lconcepts> element are the child elements of <course>.

```
<?xml version="1.0" encoding="UTF-8" ?>
- <course>
+ <information>
+ <structure>
+ <lconcepts>
</course>
```

Figure 6.15 CDL <course> tag

Information of the course will be captured through the form shown in Figure 6.16. The information will be stored in the <information> element. The <information> element is the parent of the <abtInfo> and <learningInfo> element as shown in Figure 6.17. The <abtInfo> element embeds administrative-related information about the course and includes information such as the course name (<cName>), the identifier (<cID>), the author(s) (<cAuthors>), date created (<cDateCreated>), cost of accessing the course (<cCost>), storage location (<cLocation>), and the learning domain (<cDomain>). The <learningInfo> tag embeds learning-related information about the course and includes information such as the learning preferences adopted (<cLearningPref>), the complexity

level (<cExpertise>), the level of details (<cDetail>), the hierarchical level (<cLevel>), the academic level (<cAcademic>) and the duration of the course (<cDuration>).

Figure 6.16 Screen Capture of Create New Course Form (1)

```
<?xml version="1.0" encoding="UTF-8" ?>
- <course>
- <information>
- <abtInfo>
  <cName>Integration</cName>
  <cID>MTHST0001</cID>
  <cAuthor>admin</cAuthor>
  <cDateCreated>2006-04-30 21:18:05</cDateCreated>
  <cCost>48</cCost>
  <cLocation>http://www.ntu.edu.sg/MH/TOPIC/S1/MTHST0001</cLocation>
  <cDomain>MTH</cDomain>
  <cAcademic>S</cAcademic>
</abtInfo>
- <learningInfo>
  <cLearningPref>0 -1 0.5 1</cLearningPref>
  <cExpertise>Expert</cExpertise>
  <cDetail>Normal</cDetail>
  <cDuration>2:00</cDuration>
  <cLevel>Topic</cLevel>
</learningInfo>
</information>
+ <structure>
+ <lConcepts>
</course>
```

Figure 6.17 CDL <information> tag

The structure of the course will be captured through the form shown in Figure 6.18. The <structure> element is the parent of the <internal> and <external> element. The <internal> element embeds the next lower hierarchical granularity of learning resources. For example, a topic will have its internal structure made up by RLO while a RLO will have its internal structure made up by RIO. The name of the learning resources as well as its physical location is stored by the <RLO> element and <rloLocation> element respectively. The <external> element embeds the learning resource's relationships with other learning resources. Usually, the "related" learning resources will reside at the same

hierarchical level as the main learning resource. For example, in Figure 6.16, as the main learning resource, MTHST0001, resides at a topic level, all the external courses also reside at the topic level. Similar to the <internal> element, the <external> element embeds the name of the learning resource and its physical location

Figure 6.18 Screen Capture of Create New Course Form (2)

```
<?xml version="1.0" encoding="UTF-8" ?>
- <course>
+ <information>
- <structure>
- <internal>
  <RLO1>MTHPR0006</RLO1>
  <rloLocation1>http://www.ntu.edu.sg/MTH/RLO/P6/MTHPR0006</rloLocation1>
  <RLO2>MTHPR0020</RLO2>
  <rloLocation2>http://www.ntu.edu.sg/MTH/RLO/P6/MTHPR0020</rloLocation2>
  <RLO3>MTHPR0022</RLO3>
  <rloLocation3>http://www.ntu.edu.sg/MTH/RLO/P6/MTHPR0022</rloLocation3>
</internal>
- <external>
  <cAssociated1>MTHST0059</cAssociated1>
  <assLocation1>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0001</assLocation1>
- <cPreReq>
  <essential1>MTHST0004</essential1>
  <essLocation1>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0004</essLocation1>
  <essential2>MTHST0005</essential2>
  <essLocation2>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0005</essLocation2>
  <gdToKnow1>MTHST0007</gdToKnow1>
  <gtkLocation1>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0007</gtkLocation1>
</cPreReq>
- <cPostReq>
  <augment1>MTHST0011</augment1>
  <augLocation1>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0011</augLocation1>
  <augment2>MTHST0022</augment2>
  <augLocation2>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0022</augLocation2>
  <augment3>MTHST0032</augment3>
  <augLocation3>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0032</augLocation3>
  <utilize1>MTHST0010</utilize1>
  <utiLocation1>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0010</utiLocation1>
  <utilize2>MTHST0045</utilize2>
  <utiLocation2>http://www.ntu.edu.sg/MTH/TOPIC/S1/MTHST0045</utiLocation2>
</cPostReq>
</external>
</structure>
+ <Concepts>
</course>
```

Figure 6.19 CDL <structure> tag

The <|Concepts> element, contains the learning concepts and their weightage in the main course. For example, course MTHST0001 imparts the learning concept of integration by parts (30%), integration by substitution (20%) and integration techniques (50%).

Figure 6.20 Screen Capture of Create New Course Form (3)

```
<?xml version="1.0" encoding="UTF-8" ?>
- <course>
+ <information>
+ <structure>
- <|Concepts>
  <concept1>By Parts</concept1>
  <weight1>0.3</weight1>
  <concept2>Substitution</concept2>
  <weight2>0.2</weight2>
  <concept3>Techniques</concept3>
  <weight3>0.5</weight3>
</|Concepts>
</course>
```

Figure 6.21 CDL <|Concepts> tag

6.3.2.2 Semantic Search Module

Two types of search are available to the learner: basic and advanced.

6.3.2.2.1 Basic Search

For a basic search, the learner has to specify the following mandatory fields:

1. Domain (e.g. Art, Biology, Literature, Philosophy, etc.)
2. Academic Level (e.g. Primary, Secondary, Tertiary & Others)
3. Expertise Level
4. Detail Level
5. Learning Concepts & its Corresponding Weightages

Figure 6.22 Screen Capture of Basic Search

The weightages should add up to 1. An extra feature is available to aid the learner to get the total weightages filled up so far. Below is the illustration of the basic search.

Figure 6.23 Screen Capture of Basic Search with Error

With the above information, the system will search through the database for courses within the same domain and expertise. The retrieved search results will be ranked and the similarity index (*sim*) for each course is calculated based on the formula below.

$$sim(di, q) = \frac{\vec{d_i} \cdot \vec{q}}{|\vec{d_i}| \times |\vec{q}|} = \frac{\sum_{i=1}^t w_{i-j} \times w_{i-q}}{\sqrt{\sum_{i=1}^t w_{i-j}^2} \times \sqrt{\sum_{i=1}^t w_{i-q}^2}}$$

An illustration of the basic query output is presented.

Usage:

To calculate the basic query result population

Input:

Query values (in the forms of learning concepts and their corresponding weightages) entered by the learner.

Output:

Similarity index (*sim*) for each course in the search results

Example:

Learner wants to learn Algebra (weightage 0.4), (Integration) By Parts (weightage 0.2) and (Integration) Substitution (weightage 0.2), (Integration) Techniques (weightage 0.1) and Statistics (weightage 0.1) in Domain MTH (mathematics).

Population Steps:

- Form the 2-D search matrix, $Q = \{ \begin{array}{l} \text{Algebra, 0.4} \\ \text{By Parts B, 0.2} \\ \text{Substitution, 0.2} \\ \text{Techniques, 0.1} \\ \text{Statistics, 0.1} \end{array} \}$
and $\vec{q} = \{0.4, 0.2, 0.2, 0.1, 0.1\}$
- Search all learning courses from domain MTH.
- Retrieve document that contains at least one learning concept.
i.e. the system contains 50 documents but only 6 relevant documents are retrieved
Document 1 = { Algebra, 0.3 }
Document 2 = { Algebra, 0.2 }

Statistics, 0.2 }

.....

4. Given these 6 documents, the systems will calculate the *sim* value using the formula. If the concept matches, then assign a true (value of 1) to it and multiply by the weightage.

For Document 1, only one concept matches the query. Hence $\vec{d}_1 = \{0.3, 0, 0, 0, 0\}$

$$sim = \frac{(0.4, 0.2, 0.2, 0.1, 0.1) \times (0.3, 0, 0, 0, 0)}{\sqrt{0.4^2 + 0.2^2 + 0.2^2 + 0.1^2 + 0.1^2} + \sqrt{0.3^2}} \approx 0.1482$$

For Document 2, 2 concepts match the query. Therefore $\vec{d}_2 = \{0.2, 0, 0, 0, 0.2\}$

$$sim = \frac{(0.4, 0.2, 0.2, 0.1, 0.1) \times (0.2, 0, 0, 0, 0.2)}{\sqrt{0.4^2 + 0.2^2 + 0.2^2 + 0.1^2 + 0.1^2} + \sqrt{0.2^2 + 0.2^2}} \approx 0.1261$$

5. Present all the documents according to the *sim* values.

SEARCH COURSE

Simple Search (Compulsory)

Domain:

Academic Level:

Expertise Level: ☐ Novice ☐ Intermediate ☒ Expert

Detail Level: ☐ Overview ☒ Normal ☐ Detailed

TOTAL WEIGHTAGE:

Concept: Weightage:

Figure 6.24 Screen Capture of Basic Search Query

SEARCH COURSE - RESULTS FOR SIMPLE SEARCH					
Domain:	MTH - Mathematics	Concept	Weightage		
Academic Level:	Secondary	Algebra	0.4		
Expertise Level:	Expert	By Parts	0.2		
Detail Level:	Normal	Substitution	0.2		
		Techniques	0.1		
		Statistics	0.1		

Course ID	Course Name	SIM	Learning Concept & Weightage	
1 MTHST0001	Integration	0.1482	Algebra	0.3
2 MTHST0002	Advanced Algebra	0.1261	Algebra	0.2
			Statistics	0.2
3 MTHST0004	Integration Basics I	0.2649	Algebra	1
4 MTHST0005	Integration Basics II	0.1798	Algebra	0.2
			By Parts	0.8
5 MTHST0007	Integration Basics III	0.2465	Algebra	0.5
			By Parts	0.5
6 MTHST0060	Integration	0.1332	By Parts	0.3
			Substitution	0.2
			Techniques	0.5

Figure 6.25 Screen Capture of Basic Search Result

6.3.2.2.2 Advanced Search

For an advanced search, besides the mandatory fields from a basic search, the learner can also specify the following fields: Cross Domain, Cross Expertise, Cross Details and Search Levels. These fields are used to enhance the search. They act upon the results that were retrieved by the basic search.

1. Cross Domain

Learner can specify whether the search algorithm should search for courses in the same or different domain.

2. Cross Expertise

Learner can specify whether the search algorithm should search for courses in the same or different expertise.

3. Cross Details

Learner can specify whether the search algorithm should search for courses in the same or different details level.

4. Search Levels

The search level states the degree to extend the search and is a mandatory field for concept mapping. It acts upon the results obtained from the basic search and extends the search to include the retrieved course's external structure. Each group of external structure constitutes a level. For example, a search level of 2 will mean that the search will extend from the course → course's external structure (level 1) → level 1 courses' external structure (level 2). A total of 5 levels are permitted.

SEARCH COURSE

Simple Search (Compulsory)

Domain:

Academic Level:

Expertise Level: ☐ Novice ☐ Intermediate ☒ Expert

Detail Level: ☐ Overview ☒ Normal ☐ Detailed

TOTAL WEIGHTAGE:

Concept:	Weightage:	
Algebra	0.4	<input type="button" value="Del"/>
By Parts	0.2	<input type="button" value="Del"/>
Substitution	0.2	<input type="button" value="Del"/>
Techniques	0.1	<input type="button" value="Del"/>
Statistics	0.1	<input type="button" value="Del"/>

Advanced Search (Optional)

Cross Domain: ☒ Yes ☐ No

Cross Expertise: ☒ Yes ☐ No

Cross Detail: ☒ Yes ☐ No

Search Level:

Figure 6.26 Screen Capture of Advanced Search

SEARCH COURSE - RESULTS FOR ADVANCED SEARCH

Domain:

MTH - Mathematics

Academic Level:

Secondary

Expertise Level:

Expert

Detail Level:

Normal

Concept	Weightage
Algebra	0.4
By Parts	0.2
Substitution	0.2
Techniques	0.1
Statistics	0.1

SIM: 0.2713

RANK: 0.0103

	Course ID	Concept	Weightage	Level	Expertise	Detail	Relationship
1	MTHST0001	Algebra	0.3	0	Expert	Normal	-
2	MTHPT0007	By Parts	0.6	1	Intermediate	Normal	MTHST0001 - Good To Know
3	MTHPT0007	Statistics	0.4	1	Intermediate	Normal	MTHST0001 - Good To Know
4	MTHPR0020	Substitution	0.4	2	Expert		MTHPT0007 - Augment
5	GESPR0021	Techniques	0.3	3			MTHPR0020 - Utilize

SIM: 0.1261

RANK: N.A.

	Course ID	Concept	Weightage	Level	Expertise	Detail	Relationship
1	MTHST0002	Algebra	0.2	0	Expert	Normal	-
2	MTHST0002	Statistics	0.2	0	Expert	Normal	-

Figure 6.27 Screen Capture of Advanced Search Results

An illustration of the advanced query output is presented.

Usage:

To calculate the advanced query result population

Input:

Query values (in the forms of learning concepts and their corresponding weightages) and optional search parameters entered by the learner.

Output:

Similarity index (*sim*) for a group of courses in the search results

Example:

Learner wants to learn Algebra (weightage 0.4), (Integration) By Parts (weightage 0.2) and (Integration) Substitution (weightage 0.2), (Integration) Techniques (weightage 0.1) and Statistics (weightage 0.1) in Domain MTH (mathematics).

Population Steps:

- Form the 2-D search matrix, $Q = \{ \begin{array}{l} \text{Algebra, 0.4} \\ \text{By Parts B, 0.2} \\ \text{Substitution, 0.2} \\ \text{Techniques, 0.1} \\ \text{Statistics, 0.1} \end{array} \}$
and $\vec{q} = \{0.4, 0.2, 0.2, 0.1, 0.1\}$
- Check for cross-domain, cross-details or cross-expertise tolerance
i.e. given allow cross-domain, allow cross-details and allow cross-expertise
- Search all learning courses from domain MTH
- Retrieve document that teaches at least one learning concept
This is similar to basic search except for the inclusion of three additional search parameters.
i.e. Using the previous example that the system contains 50 documents but only 6 relevant documents are retrieved.
Document 1 = $\{ \text{Algebra, 0.3} \}$
Document 2 = $\{ \text{Algebra, 0.2} \}$
Statistics, 0.2 $\}$

-
5. Check the requested search level
 6. Perform advanced course retrieval

The advanced search will start from the first document (i.e. $\vec{d}_1$) and proceed through all the search results. Since $\vec{d}_1$ contains only the Algebra Concept, the search algorithm will look the course's external courses for the remaining learning concepts. Say the course has 2 good to know courses, 2 essential courses, 2 augment courses, 2 utilize courses and 3 associated courses. A total of 11 courses will form level 1 of the search.

Given that out of these 11 courses, only one course (MTHPT0007) teaches the learning concept of “By Parts” and “Substitution”. Hence, if the learner has specified “1” for the search level, the search will stop here but another two concepts are still missing from the search.

However, if the learner has specified “2” instead for the search level, the search will continue to look for the remaining two concepts. Therefore the next level (level 2) will be formed by all external courses of course MTHPT0007.

In summary, given a search level of 3, the following courses will be retrieved.

Course MTHST0001: Algebra: 0.3. (from level 0)

Course MTHPT0007: By Parts: 0.6, Statistics: 0.4 (from level 1)

Course MTHPR0020: Substitution: 0.4 (from level 2)

Course GESPR0007: Techniques: 0.3 (from level 3)

7. The new *sim* value will be calculated.
8. A ranking value will be calculated follows:
Based on the formula:

$$rank = \sum_{i=1}^n (found_i \times weighting_i \times \Delta domain_i \times \Delta expertise_i \times \Delta details_i) \times \frac{1}{\sum course \times \sum level}$$

If there are two courses that offer the same learning concept, the course will be chosen according to the following criteria:

1. level (smaller level preferred)
2. weightage (higher weightage preferred)
3. expertise (no cross-expertise preferred)
4. details (no cross-details preferred)
5. domain (no cross-domain preferred)
9. The search results are displayed.

6.3.2.3 Dynamic Concept Map Module

This module provides the tools and platform for visualizing the learning resources in the form of a concept map. Each learning resource, regardless of its hierarchical level, is presented as a node and any two nodes can be linked by a proposition that states some meaningful relationship between the nodes. Each concept map must have a master node and depending on the request of the system/learner, all other related learning resources can be mapped outwards from the master node.

6.3.2.3.1 Concept Map Format (SVG)

The graphical presentation of the learning resources such as the course expertise (E-Map), the cognitive structure (C-Map) and the personalized learning routes (P-Map) is implemented using Scalable Vector Graphics (SVG). SVG is a platform for two-dimensional graphics and is built upon other standards such as XML, JPEG and PNG for image formats, DOM for scripting and interactivity, SMIL for animation and CSS for styling. It is interoperable and supports scripting through languages such as ECMAScript. It also has comprehensive support for animation.

Figure 6.28 Concept Map in SVG format

Based on the course description (in XML format), the system can automatically convert the course description to the graphical format (in SVG format). Using the previous example (Figure 6.19 in particular), the graphical map of the course is shown in Figure 6.28. The master node is highlighted in red while the rest of the related learning resources are linked to the master node through the five governing relationships.

The Concept Map that is presented using SVG is not a static map. Information (i.e. the learning course name and its learning concepts together with its associated weightage) as well as some animation effects (i.e. font size changes progressive) pertaining to each node will be displayed when the mouse hovers over the node. Clicking on any particular node will lead to a re-direction to the associated learning resource's main learning page and/or internal learning structure.

Figure 6.29 Hovering Effect

6.3.2.3.2 Concept Map Editor (.cmap)

Besides the direct translation of the concept map from XML to SVG format, we also implement a concept map editor to enable the SMEs to manually perform concept map creating and editing. The course concept map editor environment is shown in Figure 6.30.

Figure 6.30 Course Concept Map Editor Environment

The course concept map editor consists of four panes. The Overview Pane shows a down-sized version of the concept map; the Concept Map Pane allows the graphical editing of concept map; the Tree View Pane shows the structure of the concept map; the XML Pane displays an XML format of the concept map. The Map editor environment is shown in Figure 6.25. The environment allows basic editing functions such as copy, cut, paste, redo, undo, delete as well as more advance editing functions such as group, ungroup, regroup, zoom in, zoom out, bring-to-front, send-to-back, bring forward, send backward, select all learning resources and select all relationships. Basic administrative functions such as new, open, save and print are also implemented. The editor allows the drawing of new learning resources as well as the tagging of the relationships (i.e. associated, essential, nice-to-know, utilize and augment). Concept maps drawn using this editor environment will be stored in a .cmap format (author defined) and can be exported to XML format for storage.

Figure 6.31 Editor Environment

6.3.2.4 Learner Profile Management Module

This module contains explicitly modeled assumptions that represent the characteristics of the learners which are pertinent to the system. The learning system can consult the Learner Profile Model to adapt the performance of the system according to the learner's needs (in terms of learning content, learner's preference, etc.). Learner modeling allows the system to personalize the interaction between the learners and the learning content.

The system models three dimensions of the learner: (1) Profile Dimension, (2) Cognitive Dimension, and (3) Learning Preferences Dimension.

6.3.2.4.1 Personal Profile Dimension

REGISTRATION FORM

Login Information

Username:

Password:

Reconfirm Password:

Personal Information

First Name:

Last Name:

Date Of Birth:

Nationality:

Address:

Contact No:

Email:

Next

Figure 6.32 Learner Registration Form

The learner's personal profile dimension stores the personal information about the learner. Each new learner in the system will be asked to fill in a registration form as shown in Figure 6.32. The personal profile information will be used to uniquely identify the learner.

6.3.2.4.2 Cognitive Dimension

The learner's cognitive dimension profile stores the technical expertise information of the learner. The technical expertise information will be stored as a C-Map and will be filled initially according to the expertise population survey. Subsequent expertise updating will be done through the invocation of CMPA1 or CMPA2 (see Chapter four, section 4.3.1).

6.3.2.4.3 Learning Preferences Dimension

Learning preferences are modeled using (1) preferred learning styles, (2) preferred learning settings, and (3) learning constraints.

6.3.2.4.3.1 Preferred Learning Styles Modeling

Learning style research is drawn out of studies about the psychological, social, and physiological dimensions of the educational process which are complex. To the best of the author's knowledge, it is yet to be precisely defined. Currently, there are several scholarly works that provide a range of working models that help to define the mysterious terrain between the learner and their preferred learning styles. Among these, Kolb's Theory of Learning Styles [Kolb, 1984] and Richard M. Felder and Linda K. Silverman's Index of Learning Styles (ILS) [Felder, 1988] are some of the few widely adopted models. Kolb's Theory of Learning Styles although heavily critiqued [Rogers, 1996; Reynolds, 1997; Holman, 1997; Miettinen, 2000], should be acknowledged for creating a dramatic impact on the design and development of lifelong learning models. Felder and Silverman on the other hand, synthesized findings from a number of studies to formulate a learning style model with dimensions that are particularly relevant to science education [Felder, 1993].

As the system is dealing with science education, the system adopts ILS to assess the learning preferences. ILS is based on a four-dimensional (active/reflective, sensing/intuitive, visual/verbal, and sequential/global) learning style model.

Characteristics	Dimension	
	Active	Reflective
	<ul style="list-style-type: none"> Prefer doing something active with it--discussing or applying it or explaining it to others Prefer group work 	<ul style="list-style-type: none"> Prefer to think about it quietly first Prefer working alone
	Sensing	Intuitive
	<ul style="list-style-type: none"> Prefer learning facts Prefer solving problems by well-established methods and dislike complications and surprises Resent being tested on material that has not been explicitly covered in class Tend to be more patient with details and good at memorizing facts and doing hands-on (laboratory) work Tend to be more practical and careful Does not like courses that have no apparent connection to the real world 	<ul style="list-style-type: none"> Prefer discovering possibilities and relationships Prefer innovation and dislike repetition Tend to be better at grasping new concepts and are often more comfortable with abstractions and mathematical formulations Tend to work faster and be more innovative Does not like "plug-and-chug" courses that involve a lot of memorization and routine calculations
	Visual	Verbal
	<ul style="list-style-type: none"> Prefer pictures, diagrams, flow charts, time lines, films, and demonstrations 	<ul style="list-style-type: none"> Prefer written and spoken explanations
	Sequential	Global
	<ul style="list-style-type: none"> Prefer to gain understanding in linear steps, with each step following logically from the previous one Tend to follow logical stepwise paths in finding solutions 	<ul style="list-style-type: none"> Prefer to learn in large jumps, absorbing material almost randomly without seeing connections, and then suddenly "getting it" Tend to be able to solve complex problems quickly or put things together in novel ways once they have grasped the big picture, but they may have difficulty explaining how they did it

Table 6.2 ILS Model Comparison***Learning Styles Questionnaire and Scoring Procedures***

During registration, the learner has to take a 44-item learning styles questionnaire. This questionnaire, adopted from <http://www.engr.ncsu.edu/learningstyles/ilsweb.html>, is incorporated into the learning system to provide an estimate of the learner's learning style. The questionnaire requires the learner to answer every question on the ILS scoring sheet. There are 44 questions in total and 2 options, "a" or "b", to each question. All the answers given will be stored into the database. At the same time, the learner's ILS scoring is also calculated. This is achieved by the following steps:

1. To sum up the total number of "a"s and total number of "b"s selected as the answers for the 4 types of learners namely ACT/REF, SNS/INT, VIS/VRB and SEQ/GLO.
 - a. ACT/REF
 - Starts from Question 1 and next questions are at an interval of 4.

(Group 1: Question 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41)

b. SNS/INT

- Starts from Question 2 and next questions are at an interval of 4.

(Group 2: Question 2, 6, 10, 14, 18, 22, 26, 30, 34, 38, 42)

c. VIS/VRB

- Starts from Question 3 and next questions are at an interval of 4.

(Group 3: Question 3, 7, 11, 15, 19, 23, 27, 31, 35, 39, 43)

d. SEQ/GLO

- Starts from Question 4 and next questions are at an interval of 4.

(Group 4: Question 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44)

Each group of questions represents a type of learner. Using “for loop”, the answers selected, “a” and “b” are summed up separately for the respective group of questions.

The algorithm for calculating the ILS ACT/REF counting is given in Algorithm 6.1.

Usage:

To calculate ACT/REF Count

Input:

Answers from Group 1’s questions

Output:

Total number of “a”s and total number of “b”s selected as the answers

Algorithm:

1. let size = number of questions (in this case, 44)
2. let i = index of question
3. let count_ActRef_a = number of questions with “a” as the selected answer
4. let count_ActRef_b = number of questions with “b” as the selected answer
5. for i = 0; i < size; i++, loop
 - if answer = “a”,
 - count_ActRef_a ++;
 - else if answer = “b”,
 - count_ActRef_b ++;
- End if
- Next
- End for
6. end

Example:

Output: count_ActRef_a = 7 & count_ActRef_b = 4
(i.e. 7 “a”s & 4 “b”s)

Algorithm 6.1 ILS ACT/REF Counting Algorithm

2. To calculate the scores for the 4 types of learners based on the formula given below.

$$\text{Score} = (\text{Larger} - \text{Smaller}) + \text{Letter of Larger}$$

Using the same example from step 1 where 7 “a”s and 4 “b”s are obtained from the questions in Group 1 representing ACT/REF type of learners, the algorithm for calculating the ILS Score is given in given in Algorithm 6.2.

Usage:

To calculate Score

Input:

Number of “a”s and number of “b”s (Output from Step 1)

Output:

Score

Algorithm:

1. let larger = the larger number
2. let smaller = the smaller number
3. let letterLarger = letter of the larger number (“a” or “b”)
4. if count_ActRef_a > count_ActRef_b
 - larger = count_ActRef_a;
 - smaller = count_ActRef_b;
 - letterLarger = “a”;
- else if count_ActRef_a < count_ActRef_b
 - larger = count_ActRef_b;
 - smaller = count_ActRef_a;
 - letterLarger = “b”;
- End if
5. score = (larger – smaller) + letterLarger;
6. end

Example:

Input: 7a & 4b

Output: Score = (7 – 4) + “a” = 3a

Algorithm 6.2 ILS Scoring Algorithm

3. To calculate the rating based on the score retrieved in Step 2. The table below illustrated the rating given based on the score.

Rating	-1	-0.5	0	0.5	1
Score	9a to 11a	5a to 7a	1a to 3a or 1b to 3b	5b to 7b	9b to 11b

Table 6.3 ILS Scoring Chart

The ILS rating algorithm is presented in Algorithm 6.3.

Usage:

To calculate Rating

Input:

Score from Step 2

Output:

Rating based on the ILS Scoring Chart (Table 6.3)

Algorithm:

1. let absoluteRating = 0;
2. let prefix = “”;

```

3. let number = score.substring(0, score.length() - 1);
4. let letter = score.substring(score.length()-1, score.length());
5. if number >= 1 && score <=3)
    absoluteRating = 0;
  else if number >=5 && number <=7)
    absoluteRating = 0.5;
  else if number >=9 && number <=11)
    absoluteRating = 1;
  End if
6. if letter = "a" && absoluteRating != 0
    prefix = "-";
  End if
7. rating = prefix + absoluteRating;
8. end

```

Example:

Input: 3a

Output: rating = 0

Algorithm 6.3: ILS Rating Algorithm

Learning Preferences - Questionnaire

1. I understand something better after I
 - ☐ (a) try it out.
 - ☐ (b) think it through.
2. I would rather be considered
 - ☐ (a) realistic.
 - ☐ (b) innovative.
3. When I think about what I did yesterday, I am most likely to get
 - ☐ (a) a picture.
 - ☐ (b) words.
4. I tend to
 - ☐ (a) understand details of a subject but may be fuzzy about its overall structure.
 - ☐ (b) understand the overall structure but may be fuzzy about details.

Figure 6.33 Screen Capture of Questionnaire

6.3.2.4.3.2 Preferred Learning Settings

Besides learning preference catering, the system also models the learning settings that are preferred by the learner. The system implements a Learning Setting Model that combines the Stereotype User Model and the Overlap User Model.

The Stereotype User Model is the simplest way of user modeling. For each dimension of user modeling, the system creates a set of typical or ‘stereotype’ users (i.e. novice – beginner – intermediate - expert). A particular user is usually modeled by assigning this user to one of the stereotypes for each dimension of classification (i.e. intermediate for biology, novice for software engineering). A Stereotype User Model can either be represented as a set of pairs with Boolean values (i.e. true or false – user either belongs or does not belong to a particular stereotype group) or by some probabilistic value (value

represents the probability that user belongs to a particular stereotype group). A Stereotype User Model is useful when a quick but not necessarily completely accurate assessment of the user's background knowledge is required [Kobsa, 1993].

The Overlay Model, on the other hand, is widely used to present the user's knowledge of the subject based on the structural model of the subject domain. The basic idea behind this model is to represent an individual user's knowledge of the subject as an "overlay" of the domain model. For each domain model concept, an individual overlay model stores some values which are an estimation of the user knowledge level of this concept. These values are updated as the user progresses through the system. The values can be a binary value (known-not known), a qualitative measure (good-average-poor), or a quantitative measure, such as a probability that the user knows the concept. This model requires the knowledge domain to be modularized (i.e. topics or concepts) and its complexity depends on the levels of granularity.

Hence, the Learning Setting Model combines the Stereotype and Overlap techniques to provide a more comprehensive model of the learner's preferred learning setting. Using the Stereotype techniques, the system categorizes the learner initially by stereotype. As per the cognitive modeling, the learner will be categorized in terms of academic domain. Two forms of stereotype are used: Expertise Stereotype and Detail Stereotype. Expertise Stereotype consists of three levels: novice, intermediate, and expert level. Detail Stereotype consists of three levels: overview, normal, and detailed level. The initial classification is based on the most pre-dominant learning setting that is present during the initial expertise population.

After the initial stereotyping, the model is gradually modified as the Overlay techniques are applied to new information that are acquired from the student's interaction with the system. Now, instead of stereotyping at only the domain level, the system (or the user can explicit state) can adjust the learner's preferred setting at a lower granularity level (i.e. at RLO or RIO level); i.e. that is, for the same domain, different settings are allowed for different courses. Three types of approaches are used in the acquiring of information.

1. Observation of learner-content interaction

2. Information from C-Map - the course taken, the expertise and detail level chosen, and the assessment results
3. Regular questionnaires and test to revise the learner profile model

6.3.2.4.3.3 Learning Constraints

A learner can also explicitly specify the learning constraints. Learning constraints that are modeled in the system are (1) total time per learning session, (2) total cost per learning session, (3) preferred 'include' list of publisher (of learning content), and (4) preferred 'exclude' list of publisher (of learning content).

6.4 Summary

This chapter presents the technical implementation details of the developed prototype system. The proposed framework lays the foundation for the development of a full fledged learning environment and includes considerations for supporting other technological advances. As a proof of concept, a prototype system is developed and documented in this chapter to bring together the research work.

CHAPTER 7

TYPICAL LEARNING SCENARIO

The typical learning scenario and its learning problems, introduced in Chapter 2 and constantly revisited throughout this thesis, is revisited and summarized in this chapter. Through a re-enactment of the learning scenario, we put together all aspects of our research work and present a learning environment that truly provides learning personalization.

7.1 Learning Problems Rectified

The learning problems presented in Chapter 2 are revisited. Here, we summarized how our proposed theorems, methodologies and algorithms contribute to the resolution of these problems.

1. A one-size-fits-all learning solutions for all learners

One of the main aims of learning personalization is to solve the one-size-fits-all learning problem. Like Adaptive Hypermedia Systems, we also rely on the effective gathering of the learner profile (includes learning preference and constraints such as limitation on the amount of time/effort/money to be spent on each learning solution). However, we go one step further in the creation of a personalized learning solution by (1) dynamically assembling a course that starts off from the learner's prior knowledge and (2) providing different abstraction of the same learning concept. Our novel method of dynamically assembling a course that starts off from the learner's prior knowledge is explained in Chapter three and illustrated in Appendix A. The ability to provide different abstractions of the same learning concept is provided through the use of concept maps. In the learning experience, it is common to find a learner having to focus on a particular issue that is specific to his interests or capabilities. Such a learning phase can become particularly intense and difficult if the learner has not assimilated enough knowledge of the basic concepts. Mental processes that lead to the organization of learning material into visual maps seem to be particularly helpful for understanding parts of a topic, which normally would escape even a careful reader [Cicognani, 2000]. Hence, depending on the level of detail required, concept maps can be used as an overview, or at any abstract level of a knowledge domain. The level of abstraction can be based on the learner's expertise level where the more abstract logic is presented only to higher order learning or upon an explicit learner's request. Highly abstract concept maps may even be used as meta-maps linking to subordinate maps, which drill down to greater levels of granularity and specification.

2. Recommended solutions do not take the learner's learning preferences or expertise into account

Currently, learning preferences are provided for by the use of learner profiles. However, to the best of our knowledge, no learning personalized systems are able to cater for learning contents according to the expertise of the learners. Through the usage of our novel method of learning dependency and the concept of expertise gap identification (explained in Chapter four), the expertise of the learner can be presented in the form of a concept map (see Figure 7.2). As concept maps can be effectively programmed using XML (Extensible Markup Language) and displayed using SVG (Scalable Vector Graphics) and/or our novel .cmap format, its ease of representation enables it to be easily compared and synthesized with other concept maps. Hence, through such a comparison and synthesis between the learner's query and his expertise in the search query domain, we are able to generate a concept map depicting all the pre-requisites and relevant knowledge that the learner needs to acquire. In this way, we can effectively assemble a learning course that is initiated from the learner's cognitive structure.

3. Inaccurate learning solutions recommendation due to keyword match techniques

Concept maps can be described by XML. Hence, they can be easily meta-tagged with information depicting the learning concepts that are associated with each learning node in the map. Besides learning concepts, a weightage depicting the importance of the learning concept for a particular learning document can also be added. This method is far superior to keyword match that is often employed by the search engines. This is because in most cases, a particular learning document that teaches some key concepts might not even contain the keyword itself. Also, often, the important key concepts are usually not the terms that occur most frequently in the course materials (argument proven by experiment conducted – see Appendix D). Furthermore, the assignment of weightage is by the content expert. Hence, unscrupulous attempts to increase the learning document ranking can be eliminated.

4. No learning support provided

Learning support is an important aspect of learning personalization. Unfortunately, this is not realized by existing personalized systems. In this research, we utilize the Goal Model and learning methodologies to provide the learning supports.

5. Stand alone learning solution with no follow-up recommendations to other relevant or related courses

In most learning instances, the concepts or knowledge that the learner wishes to master usually does not reside in a single course or at a particular granularity level; i.e. in order to master the desired learning concept(s), the learner may have to take more than one courses and these courses might reside at different granularity level (e.g. course/topic/article level) or may even be cross-domain. Thus, in order to provide a complete learning experience that aims to cover all aspects of the learner's learning requirements, the recommended solution may need to propose a training plan that is based on more than one courses. While these courses usually exist as a stand-alone basis in the course repository, all the relevant courses can be gathered to form different nodes in a single concept map. Thus, instead of presenting the training solution as independent learning material, the training solution will consist of a series of inter-related courses. The visual image of the learning paths is presented through the usage of the concept mapping technique and labeled as the Course Expertise Map (E-Map) (see Figure 7.3).

6. Teaching plans are not transparent

Current classical planning systems for learning/teaching are often not transparent as the details about learning plan generation are hidden inside software components. This makes it difficult for pedagogists to understand, and in consequence, to trust them. The proposed goal driven learning methodology contributes to this area by considering different components in the planning and execution process and making them explicitly into goals. Through this formal framework, each goal, which represents the planning and execution process, are not only explicit but can also be easily customized according to domain knowledge and requirements. Also, the execution of the learning goals can be easily verified and validated by the domain experts through the use of checklist for increased credibility.

7.2 Scenario Revisit

The same scenario provided in chapter 2 is revisited. Now, instead of searching the Net for learning solution, the learner can utilize our system to search for learning solutions. However, instead of considering only the learner's query (as per existing textual search approach), our system will also investigate the learner's academic background to see if the learner possesses any relevant knowledge in the query domain. Moreover, besides the pure textual display of ranked learning solutions, the system will also augment the textual display with an optional graphical display of learning solution.

A typical learning request will trigger the formulation of a goal net as shown in Figure 7.1. As an illustration, the full goal net is not presented. Here, we use the partial goal net to present how learning personalization and learner support is being implemented. The focal goal of the agent is "*Master Concept*". The focal goal is further decomposed into a set of sub goals S_1 , S_2 , S_3 , S_4 , S_5 , and S_6 connected via state transitions.

Figure 7.1 Partial Goal Net Solution

- $S_1.id = 1$, $S_1.Name = \text{"Query Analyzed"}$
- $S_2.id = 2$, $S_2.Name = \text{"C-Map Retrieved"}$
- $S_3.id = 3$, $S_3.Name = \text{"E-Map Generated"}$
- $S_4.id = 4$, $S_4.Name = \text{"P-Map Generated"}$
- $S_5.id = 5$, $S_5.Name = \text{"Learning Scaffolded"}$
- $S_6.id = 6$, $S_6.Name = \text{"Final Assessment"}$

Once the learning query is submitted, Task List A will be formed. The primary aim of this task list is to analyze the search request and to retrieve the academic domain whereby the learning solution will be based on. The results of Task List A will then be fed to state S_2 and state S_3 which will output the C-Map and E-Map respectively.

Figure 7.2 Cognitive Map (C-Map)

Figure 7.3 Course Expertise Map (E-Map)

The C-Map is our novel method of eliciting the learner's prior knowledge. It is also one of the essential inputs to our learning personalization methodology. Through such an externalization of cognitive structure, we can recommend learning courses that takes the learner's prior knowledge into consideration – *an important learning factor that is neglected in existing learning personalization approach*. In this example, the learner's C-Map in the mathematics domain is retrieved.

Also, existing solutions often present the search solution in a textual form. Depending on the ranking of the keyword match, course MTHST0005 / MTHPT0003 (the concept of ratio will be taught by course MTHST0005 while the concept of area will be taught by course MTHPT0003.) will be set as the most appropriate match while the other course, although having the same priority, will be ranked lower. Hence, the learner often has to sift through the search results to identify the relevant courses to master. The usage of E-map however eliminates such problems. Besides showing all the requested courses (highlighted in red) in a single map, other relevant courses can also be displayed; that is, besides showing the requested courses (highlighted in red), the external relations of the requested courses to other courses are also shown (in yellow). For example, MTHST0005 requires 3 essential pre-requisites courses (concept of proportion II, percentage II and measure VI) to be covered first. It is also linked to a good-to-know pre-requisites concept. While the good-to-know pre-requisite concept is not essential for understanding MTHST0005, it can enhance the understanding of the course. Post-requisites and associated courses are also depicted. These related courses while not essential to master the requested concepts, can provide the student with a broader conceptual understanding. Besides, it gives the student the luxury of augmenting his understanding on the learning concepts later at his own convenience.

While most systems stop at this stage and proclaim learning personalization, we argued that such solutions still belong to the ‘one-size-fits-all’ category. Every individual, with the same set of query, will be presented with the same learning solution. In this research, our *Learning personalization* is performed by comparing the C-Map and E-Map. Through the comparison between the requested course pre-requisite and the learner’s expertise, a personalized learning path that starts from the student’s prior knowledge (blue) can be mapped. An example of a personalized map is depicted in Figure 7.4. It is important to note that instead of recommending course MTHST0005 and MTHPT0003 as per current education practice, learning is brought ‘backwards’ in the sense that the starting point of the training is traced all the way back to the student’s prior knowledge (blue). Hence, instead of taking course MTHST0005, the student should first take courses MTHPT0022, MTHPT0023, MTHPT0027 before attempting the higher-level courses.

Alternatively, the student can also revisit the courses that they have previously taken for revision purposes; i.e. MTHPT0009, MTHPT0011, MTHPT0010.

Figure 7.4 Personalized Map (P-Map)

We argue that only when we can prescribe a learning solution according to the learner's prior knowledge, can we claim to have provided learning personalization. Moreover, since no two learners have the same prior knowledge, the personalized solution that we recommend will never be the same for different learners.

Another advantage of our method lies in the provision of search query in terms of learning concepts. Figure 7.4 is a typical example where a search on the learning concept of ratio and area generates a concept map which centers on key learning concepts (highlighted in red). Other learning concepts presented in this map depict the pre-requisite learning concepts of the key concepts.

For discussion's sake, only a single learning route is presented. The actual system is able to recommend different learning routes as some concepts can be covered by more than one course offering. Also, each learning route is characterized by (1) number of courses, (2) expertise level, (3) detail level, (4) total cost, (5) total time, (6) style of content presentation and (7) dominant learning preference (the learning resources parameters have been discussed in chapter six). Hence, based on the learner's learning preferences and constraints, a variety of training routes can be recommended.

Our learning personalization offering does not stop here. We further augment our learning personalization approach by providing learning support and/or learner's goal provision. Specifically, we exploit the use of agents to mimic the guidance and support of a personalized mentor. Through appropriate assessment of the learner's capability and appreciation of the learning content at predefined intervals, different levels of student support in the form of scaffolds will be accorded. Besides learning support, the scaffolding level will also determine the type of learning content presentation, granularity of content offering, as well as the frequency of assessment. A total of three scaffolding support levels can be provided (see Chapter 3 Section 3.3.2.3.1).

The scaffolding support is performed by state S_5 . As state S_5 is a composite state, we use the second level of the goal net hierarchy to present its flow. State S_5 is decomposed by its associated goal net into state S_{51} , S_{52} , S_{53} , and S_{54} where:

$S_{51}.id = 1$, $S_{51}.Name = \text{"Content Presented"}$

$S_{52}.id = 2$, $S_{52}.Name = \text{"Assessed"}$

$S_{53}.id = 3$, $S_{53}.Name = \text{"Scaffolding Level Assigned"}$

$S_{54}.id = 4$, $S_{54}.Name = \text{"Content Customized"}$

Using the courses selected by the learner from the P-Map (Figure 7.4), the content presentation will be initially formulated based on his learning preferences (State S_{51}). In the course of learning, the learner will be continuously assessed at regular interval (frequency of assessment will be determined at real-time according to the scaffolding table) and student support in the form of scaffolds will be accorded according to performance. The content presentation will also be changed accordingly.

Besides providing learning personalization and learning support, the goal net model can also be used to conceptualize any learning goals at real time. In addition, as each state has its own ability to dynamically assess the existing situation and calculate its worth value (in terms of dynamic assessment of situation changes, satisfaction gained and completeness value), the satisfaction gained from the goal achievement can be quantitatively measured and compared. Hence, this effectively enables the agent to

evaluate, report its progress and intelligently choose the next sub goal autonomously for reaching the final goal.

7.3 Contributions

Briefly, through our learning personalization methodology and Goal Model, we have shown that our approach is superior to existing approaches through the following:

1. **Provision of learning content sequencing and structure that starts off from the learner's prior knowledge.** This is in line with the cognitivism and constructivism theorems that require learning to be founded on the premise that learners learn by reflecting and constructing their own understanding based on their prior experiences and knowledge. Like the two theorems, our content sequencing methodology is based on the emphasis that memory plays a prominent role in the learning process.
2. **Provision of scaffolding through the use of fading and student support.** This is in line with Lev Vygotsky's sociocultural theory and his concept of the Zone of Proximal Development (ZPD).
3. **Provision of dynamic content presentation format (problem-based / case-based / scenario-based) according to assessment and scaffolding evaluation.** This is in line with the behaviorism theorem that the arrangement of stimuli and consequences form the most critical factor in learning. Through the use of different learning situation, the learner's behavior is shaped.
4. **Provision of learner's goal specification and fulfillment.** This is in line with the learner-centric approach of learning. Here, the learner's goals take centre stage and is the dominant factor in the learning process.
5. **Provision of a dynamic concept map-based learning content structure.** The employment of concept maps aims to nurture self-directed learning [Farrand, 2002], create a greater awareness of learning responsibility and make learning meaningful. It also incorporates the significant roles of motivation and volition in initiating and maintaining the learner's effort and focus [Garrison, 1997].
6. **Provision of addition search options using learning concept.** This method complements the existing approach of keyword match. However, in order to take full

advantage of this search option, learners must have some basic knowledge of the learning concept requested.

7. **Provision of a complete learning solution through a series of inter-related courses.** This method presents a more complete learning solution. Also, it provides an avenue and scaffold for learners to either augment or utilize their newly acquired knowledge.

7.4 Summary

As educators, we hone our craft by continuously reflecting on how to leverage the best of technological advancements to satisfy the demands of the learners as well as to cater to the process by which learning takes root. This inquiry persists in the pursuit for learning personalization. Through our review of existing literature, we unveil disturbing trends that exist in learning personalization approaches, which while able to provide relevant learning contents, are still ‘learning impaired’. Hence, we have initiated a practical approach that attempts to fill the gap between what the learners need and what we can currently provide. Specifically, we present, through a learning scenario, how the problems of learning personalization and student support can be addressed with our research approach. Besides learning personalization, our goal-oriented modeling approach can also be used in the area of planning. Through the notion of goals, a comprehensive approach for specifying and validating the plans of the learner’s learning objectives can be achieved.

CHAPTER 8

CONCLUSION AND FUTURE WORK

This chapter concludes the field of research and presents the future work.

8.1 Conclusion

New technologies, particularly information and communications technology, coupled with the impact of the knowledge economy has put the provision of what constitutes satisfactory education and training under pressure. Educators are now facing the same challenges that the business community has faced. These include the loss of market share, increased competition and increased global demands for productivity and quality with diminished resources [Watkins, 2005]. In view of how technology has for decades played a huge role in revitalizing business, many educational institutions have turned their focus to technology and hope it can offer the same revitalizing effect.

However, the tools technology place at our disposal should not define the tasks of education nor should technology shape the student users, it has, together with the changing landscape of education and the changing mindset of the students, upped the ante for the educational sector and forms a critical pressure point for challenging the dominant assumptions and characteristics of existing traditionally organized institutions of learning. The proliferation of online solution providers and the challenges posed by distance learning are also issuing a serious ‘wake up call’ to the educational sector. Even though the distance learning vision for the future is not yet realized, it is clear that many of its proposed conceptual philosophies have gained currency in the educational sector [Parker, 2004]. These conceptual philosophies include a shift from episodic curriculum learning to clusters of instructional resources to integrated perpetual learning. The role of educators has also been re-conceptualized – from content dictators to learning mentors and facilitators. The most pervasive of these changes is the transition of instructions from teacher directed to learner directed learning.

Unfortunately, while all these changes are important and pushing the world for adoption of a new educational model to share and reuse information created by various individuals and organizations, much needs to be done in the receiving end of this important initiative – the students and the instructors. Current educational systems are still unable to provide an educational environment that honors and recognizes the unique gifts, skills, passions and attributes that characterizes each learner. In spite of the overwhelming research to the

contrary, the assumption that a ‘one-size-fits-all’ traditional learning model can and shall effectively serve all students, continues to plague the educational sector. As a result of this, many educational setups has been reduced to merely a depository portal – instructors uploading their existing Word, PDF or powerpoint files for students to download. Thus, the real potential of using information technology for learning and teaching has not been unleashed.

It is time to acknowledge that our traditional model can no longer effectively serves the needs of all learners. We believe that instead of reinventing the traditional model or trying to force-fit every learner into a model, we can add to a model that offers the flexibility of personalized learning as well as fulfill the educational philosophy of self-advocacy, self-determination and self-directed learning.

In this research, we present a learning personalization and content development methodology that is rooted in deep pedagogical principles. Through the content development methodology, we revise the ADDIE model and use it as an equivalent process in education. The content development methodology lays the foundation for vital e-learning requirements such as learner-centric, concept-map based and personalized aspects of learning to be implemented. More importantly, essential pedagogies and scaffolding strategies are devised and integrated into the methodology. The learning personalization methodology provides the theorems, algorithms and mathematical models that underpin this research. An important contribution of the learning personalization methodology is the novel method of externalizing the learner’s existing cognitive structure using the concept mapping approach. Lastly, a goal driven learning model is used as an implementation platform. Besides catering for personalized learning, goals can also be used to model and assist training or the learning preparation processes.

8.2 Contributions

The main contributions of this research include:

1. **A summary of the state-of-art of e-learning status and systems especially focusing on pedagogy and learning personalization** (Chapter 2)

2. A novel pedagogical content development methodology

There have been disparate efforts on characterizing the content for e-learning but little work has been done to integrate pedagogical considerations as part of the content development process. Most e-learning systems often operate under the constraints of a limited theoretical background or rely on some general theories of learning. The translation of behaviorism, cognitivism, and constructivism theorems – from an era where learning is not impacted by technology to one where technology is deeply embedded into learning – in the e-learning context, is one major research contribution. (Chapter 3)

3. A novel 4-T model

The 4-T model presents a systematic translation and identification of e-learning activities based on the psychological principles of learning. (Chapter 3)

4. A novel learning personalization methodology

The learning personalization methodology enables the learner's cognitive structure to be externalized in a concept map format. This novel method, based on deep cognitivism theorems, allows the learning resources to be personalized and start off from the learner's prior knowledge region. (Chapter 4)

5. A goal driven learning model

A learning program that uses goals to guide its understanding and appreciation of information and context is a great improvement over one that processes everything in detail. We believe that our goal driven learning system is far superior to existing learning systems which are data-driven or object-driven. The added information (i.e. priority, focus, contingency plans) and the nature of goal driven learning (active, opportunistic, multi-strategy and realistic) allows a better learning solution as learning is simplified to become a goal-directed transformation of knowledge. (Chapter 5)

6. A mathematical model for partial goal measurement

Traditional goal-oriented approaches to building intelligent agents only consider the absolute achievement of a goal (either achieved or not achieved). However, in a learning domain that is dynamic and continuous, the absolute achievement of a goal may not be always feasible or even possible. In most instances, most goals can only

be satisfied partially. Hence, in order to model a learning domain accurately, such situations must be modeled. Also, the quantification of goals provide a more comprehensive learning solution as the system can assess the goal measurement value to improve the learning solution (i.e. activate contingency plan if goal measurement value falls below certain value). (Chapter 5)

7. Practical and realistic

Case studies and scenarios have been conducted to explore the usage of the proposed methodologies and models in various learning application domains. The results of these case studies and scenarios show that the proposed methodologies and models are not only promising but practical. (Chapter 5, Chapter 6 and Chapter 7)

In conclusion, this research contributes to the e-learning state of art by introducing a learning personalization model that explores new perspectives on personalizing Web learning instructions and environment. Recognizing that conventional classroom designs may not work for e-learning (especially true when an instructor is not around to stimulate motivation and continual learning progress), sufficient pedagogical considerations and learning support in the form of scaffolds is proposed. A Goal-Driven Learning paradigm is developed to support the proposed theories and to facilitate the design and implementation of the learning environment.

8.3 Future Work

In spite of the reported contributions mentioned above, the current work can be extended in at least the following dimensions:

1. Further exploration of learning theories and learning outcomes to enhance the pedagogical aspects of the content development methodology
2. Further exploration of the goal driven learning model for modeling mobile agents
3. Extension of the goal driven learning model to include negotiation between agents
4. Further development of the prototype system to extend it to a full-fledged learning environment

8.3.1 Service-Oriented Learning Resources Structure

E-learning is centered on the issue of reusability and the concept of RLO. Unfortunately, while the concept of usability appears to be an attractive proposition to yield a justifiable return on investment, the physical implementation of the RLO concept is difficult. Today, there are no known e-learning systems that are completely reusable. Although there are bits and pieces of content here and there within the LCMS that can be reused, rarely is everything reusable. Rather, most of current RLOs are repurposed rather than reused. There is a distinction to be drawn between reusability and repurposement. Reusability is described as the ability to take the whole learning object as it is and to reuse it wholesale, whereas repurposement refers to the ability to extract certain portions of the learning objects, modify part of it, and readapt to another new learning context. Currently, most if not all RLO are being repurposed, as the learning object is not always extensible beyond the context in which it was produced. While content repurposing allows the learning objects to become customizable and thereby promotes reuse, it is not the main intent for which the term RLO is coined.

With the market demanding a quicker and less-expensive way to build and maintain content, research into creating a truly reusable learning object is becoming an imperative. Other than RLOs, there are no other development strategies that have emerged that promise a quicker time to market, reduced cost to produce learning material, and a single maintenance source for whatever courseware that needs updating.

Here, we propose extending the research to develop a feasible and efficient way to structure the RLOs or learning resources so that they can be easily reused. Besides reuse, the learning resources must also be delivered using a binary map. The learning resource is first encapsulated according to a pre-defined content structure that is governed by the pedagogical content development model. The content is then further enveloped into binary maps and cast as a service set in a Service-Oriented Architecture (SOA). Based on the Goal Driven learning Model and the creation of different learning goals, the search engine must be able to match a goal to a special combination of RLOs to create a personalized binary map. Different taxonomy and metadata strategies for effective content management to address the use and reuse of learning content in the e-learning system must also be developed.

Publications

Journal and Transaction Articles

1. **Chao Boon Teo**, Robert Kheng Leng Gay, “Knowledge-Driven Model to Personalize E-learning” ACM Journal on Educational Resources in Computing (JERIC), Volume 6, Issue 1 (March 2006), ISSN:1531-4278.
2. **Chao Boon Teo**, Habib Mir M. Hosseini, Zhiqi Shen, Robert Kheng Leng Gay, “Goal Oriented Learning Personalization”. Journal of Research in Science and Technology Education. (To appear)
3. **Chao Boon Teo**, Robert Kheng Leng Gay, “Concept Map Provision for E-learning”. International Journal of Instructional Technology and Distance Learning. Vol. 3 No. 7. ISSN: 1550-6908. 2006. pp. 17 – 32.
4. **Chao Boon Teo**, Robert Kheng Leng Gay, “E-learners, A Special Breed of Learners with Implicit Learning Characteristics”. Special Issue of the International Journal of the Computer, the Internet and Management, Vol. 14 No. SP1, August 2006.
5. **Chao Boon Teo**, Robert Kheng Leng Gay, “Provision of Self-Directed Learning Using Concept Mapping”. WSEAS Transactions on Advances in Engineering Education, Vol. 3, Issue 6, ISSN 1790-1979, pp. 491 – 498, June 2006.
6. **Chao Boon Teo**, Agnes Shook Cheong Chang, Robert Kheng Leng Gay, “Pedagogy Considerations for E-learning”. International Journal of Instructional Technology and Distance Learning. Vol. 3 No. 5. ISSN: 1550-6908. 2006. pp. 3 – 26. *Selected as the Lead Article in the May, 2006 Issue.*
7. **Chao Boon Teo**, Robert Kheng Leng Gay, “Personalization Issues in E-Learning,” WSEAS Transactions on Information Science and Applications, Vol. 2, Issue 10, ISSN 1790-0832, pp. 1514 – 1522, October 2005.
8. **Chao Boon Teo**, Robert Kheng Leng Gay, “Concept-Based System Design to Personalize E-learning,” WSEAS Transactions on Information Science and Applications, Vol. 1, Issue 5, pp. 1248 – 1255, November 2004.

Magazine Articles

9. Fook Hoong Choo, Robert Kheng Leng Gay, **Chao Boon Teo**, “An Instructor-Student Centric E-learning Management System” Digital Learning Magazine, Vol. 2, Issue 7, July, 2006.

Conference Articles

10. **Chao Boon Teo**, Robert Kheng Leng Gay, “A Framework for a 21st Century E-learning System”. Proc. 7th WSEAS Int. Conf. on Distance Learning and Web Engineering (DIWEB '07), Beijing, China, Sept. 15th-17th, 2007. pp 297-303.
11. Fook Hoong Choo, Robert Kheng Leng Gay, H Mir Mohammad Hosseini, **Chao Boon Teo**, “E-learning Grid - An Online Learning Network”. Proc. 7th WSEAS Int. Conf. on Distance Learning and Web Engineering (DIWEB '07), Beijing, China, Sept. 15th-17th, 2007. pp 310-314.
12. **Chao Boon Teo**, Robert Kheng Leng Gay, “E-learners, A Special Breed of Learners with Implicit Learning Characteristics” Proceeding of The Third International Conference on eLearning for knowledge based Society. Bangkok, Thailand, August 03-04, 2006.
13. **Chao Boon Teo**, Robert Kheng Leng Gay, “Engineering Education Reforms Using Concept Mapping” Proceeding of The 3rd WSEAS/IASME International Conference on ENGINEERING EDUCATION (EE'06). Vouliagmeni, Athens, Greece, July 11-13, 2006.
14. Fook Hoong Choo, Robert Kheng Leng Gay, **Chao Boon Teo**, “An Instructor-Student Centric E-learning Management System” Proceedings of Digital Learning Asia 2006 Conference, Bangkok, Thailand, 26-28 April, 2006. *Selected for publication in the Digital Learning Magazine.*
15. Robert Kheng Leng Gay, **Chao Boon Teo**, “Redefining E-learning” Proceedings of Digital Learning Asia 2006 Conference, Bangkok, Thailand, 26-28 April, 2006.
16. **Chao Boon Teo**, Robert Kheng Leng Gay, “Scaffolding Considerations for E-learning” Proceedings of Digital Learning Asia 2006 Conference, Bangkok, Thailand, 26-28 April, 2006.
17. **Chao Boon Teo**, Robert Kheng Leng Gay, “Content Authoring System to Personalize E-Learning,” Proceedings of the 5th WSEAS Int. Conf. on

DISTANCE LEARNING AND WEB ENGINEERING, Corfu, Greece, August 23-25, 2005 (pp105-110).

18. **Chao Boon Teo**, Robert Kheng Leng Gay, “Concept-Based System Design to Personalize E-learning,” Proceedings of the WSEAS Multiconference on E-ACTIVITIES, ICECS 2004, Crete, Greece, October 2004. *Selected as Top Articles.*
19. **Chao Boon Teo**, Robert, Kheng Leng Gay, “Concept Map Approach to E-learning,” World Conference on E-Learning in Corp., Govt., Health, & Higher Ed., Vol. 2004, Issue. 1, 2004, pp. 2160-2165, Norfolk, VA: AACE.

Submitted Articles

20. **Chao Boon Teo**, Robert Kheng Leng Gay, “A Formal Goal Model for Modeling E-learning Curriculum”. Submitted to Journal of Information Technology Education (JITE).
21. **Chao Boon Teo**, Robert Kheng Leng Gay, “A Novel Content Development Methodology to Personalize E-learning”. Submitted to Journal of Information Technology Education (JITE).

References

- [Abrami, 1996] Abrami, P.C., and Bures, E.M. (1996). Computer-supported collaborative learning and distance education. *American Journal of Distance Education*, 10, pp. 37-42.
- [Abramowicz, 2003] Abramowicz, W., Kowalkiewicz M., and Zawadzki, P. (2003). Towards User Centric E-Learning Systems, in *Web Services, E-Business, and the Semantic Web*, Lecture Notes in Computer Science 2512, Springer Verlag, 2003, 109–120.
- [Alexander, 1998] Alexander, S., and McKenzie, J (1998). An Evaluation of Information Technology Projects for University Learning. Committee for University Teaching and Staff Development and the Department of Employment, Education, Training and Youth Affairs, Canberra.
- [Álvarez, 2003] Álvarez, N., and Trefftz, H. (2003). Agents as Tutors in Augmented Reality. In *proceedings of IASTED Computer Science and Technology (CST 2003)*, pp 249 - 252, Cancún, México, May 19-21, 2003.
- [Angehrn, 2001] Angehrn, A., Nabeth, T., and Roda, C. (2001). Towards personalised, socially aware and active e-learning systems. CALT White Paper.
- [Angluin, 1998] Angluin, D. (1988). Queries and concept learning. *Machine Learning*, 2, 319- 342.
- [Antón, 1997] Anton, A. I. (1997). Goal Identification and Refinement in the Specification of Software-Based Information Systems, Ph.D. thesis, Department of Computer Science, Georgia Institute of Technology, June 1997.
- [Arapi, 2007] Arapi, P., Moumoutzis, N., Mylonakis, M., and Christodoulakis, S. (2007). A Pedagogy-driven Personalization Framework to Support Adaptive Learning Experiences. *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)* pp. 96 – 97.
- [Arasu, 2001] Arasu, A., Cho, J., Garcia-Molina, H., Paepcke, A. and Raghavan, S. (2001). Searching the Web. *ACM Transactions on Internet Technology*, 1(1), 97-101.
- [Axelsson, 1999] Axelsson, J. (1999). Holistic Object-Oriented Modeling of Distributed Automotive Real-Time Control Applications. In *Proc. of the 2nd IEEE International Symposium on Object-Oriented Real-Time Distributed Computing (ISORC'99)*, pp. 85–92, Saint-Malo, France, May 1999. IEEE Computer Society.
- [Ausubel, 1965] Ausubel, D.P. (1965). A Cognitive Structure View of Word Meaning. In R. Anderson and D. Ausubel (Eds.), *Readings in the Psychology of Cognition*. New York: Holt, Rinehart and Winston.

- [Avouris, 2001] Avouris, N. and Solomos, K. (2001). User interaction with Web-based agents for distance learning. *Journal of Research and Practice in Information Technology*, 33(1), 16-29.
- [Ayala, 2003] Ayala, G. (2003) 'Towards Lifelong Learning Environments: Agents supporting the collaborative construction of knowledge in virtual communities', in Wasson, B., Ludvigsen, S. and Hoppe, U. (eds), *Designing for Change*, 141-149.
- [Ayersman, 1995] Ayersman D.J., and Minden A. (1995). Individual differences, computers and instruction. *Computers in human behavior*, 11(3-4), 371-390.
- [Baecker, 2004] Baecker, R., Wolf, P., and Rankin, K. (2004). The ePresence Interactive Webcasting and Archiving System: Technology Overview and Current Research Issues. *Proceedings ELearn 2004*, Washington D.C.
- [Baeza-Yates, 1999] Baeza-Yates, R., and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley Longman Publishing Co. Inc.
- [Balabanovic, 1997] Balabanovic, M. and Shoham, Y. (1997). Fab: Content-based, Collaborative Recommendation. *Communications of the ACM*, 40(3), 66-72.
- [Barr, 1995] Barr, R.B. and Tagg, J. (1995). From teaching to learning: A new paradigm for undergraduate education. *Change*, 27 (6), 13-25.
- [Baruque, 2003] Baruque L., and Melo R. (2003). The Various Stages of an Instructional Systems Development Methodology for e-learning Modules, *Proceedings of the I PGLDB Conference*, April 2003.
- [Basu, 2001] Basu, C, Hirsh, H., Cohen, W. and Nevill-Manning, C. (2001) Technical paper recommendations: a study in combining multiple information sources. *Journal of Artificial Intelligence Research*, 1, 231-252.
- [Baum, 1991] Baum, E., and Lang, K. (1991). Neural network algorithms that learn in polynomial time from examples and queries. *IEEE Trans. Neural Networks*, 2.
- [Baylor, 2002] Baylor, A. (2002). Agent-based learning environments for investigating teaching and learning. *Journal of Educational Computing Research*, 26 (3), 249-270.
- [Becta, 2002] Becta (2002) *Young People's Use of ICT* (Coventry, Becta).
- [Bednar, 1991] Bednar, A. K., Cunningham, D., Duffy, T. M., and Perry, J. D. (1991). Theory into practice: How do we link? In G. J. Anglin (Ed.), *Instructional technology: Past, present, and future*. Englewood, CO: Libraries Unlimited.

-
- [Bennett, 1999] Bennett, S., Priest, A. and Macpherson, C. (1999). Learning about online learning: An approach to staff development for university teachers. *Australian Journal of Educational Technology*, 15(3), pp. 207-221.
- [Bergen, 2003] Bergen, M., Denzinger, J. and Kidney, J. (2003). Teaching Cooperation in Multi-Agent Systems with the help of the ARES System, Proc, WCCCE-03, Courtenay, 2003.
- [Berghel, 1997] Berghel, H. (1997). Cyberspace 2000: Dealing with Information overload. *Communications of the ACM*, 40(2), 19-24.
- [Berk, 1995] Berk, L. E., and Winsler, A. (1995). Scaffolding children's learning: Vygotsky and early childhood education. NAEYC Research and Practice Series, 7. Washington, DC: National Association for the Education of Young Children.
- [Bruce, 2003] Bruce, B. (2003). eLearning with Dreamweaver MX: Building Online Learning Applications. Macromedia Press. pp11-31.
- [Bixler, 2000] Bixler, B., and Spotts, J. (2000). Screen design and levels of interactivity in web-based training: <http://www.clat.psu.edu/homes/jds/john/research/ivla1998/ivla98.htm>
- [Blanton, 1998] Blanton, Betty B. (1998). The Application of the Cognitive Learning Theory to Instructional Design. *International Journal of Instructional Media*, 25, 2, 171-177.
- [Bodrova, 1996] Bodrova, E., and Leong, D. J. (1996). Tools of the mind: The Vygotskian approach to early childhood education. Englewood Cliffs, NJ: Merrill/Prentice Hall.
- [Bollacker, 1999] Bollacker, K.D., Lawrence, S. and Giles, C.L. (1999). A system for automatic personalized tracking of scientific literature on the web. In Proc. ACM Conference on Digital Libraries (DL 1999), 105-113.
- [Booch, 1994] Booch, G. (1994). Object-Oriented Analysis and Design with Applications. Addison Wesley, 1994.
- [Borchers, 1998] Borchers, A., Herlocker, J., Konstanand, J. and Riedl, J. (1998). Ganging Up on Information Overload. *Computer*, 31(4), 106-108.
- [Boticario, 2000] Boticario, J.G., Gaudioso, E. and Hernandez, F. (2000) Adaptive navigation support and adaptive collaboration support in WebDL', Proceedings of the International Conference on Adaptive Hypermedia and Adaptive Web-based Systems, number 1892 in *Lecture Notes in Computer Science (LNCS)*, pages 51-61, Trento, Italy, August 2000. Springer Verlag.

- [Boud, 1992] Boud, D.J. (1992). Individualized learning systems: Individualized instruction in higher education. In M. Eraut (Ed.), *The international encyclopedia of educational technology*. New York: Pergamon Press.
- [Bounds] Bounds, G., Compliance regulations create a software niche, *Wall Street Journal Online*. Retrieved January 20, 2005 from <http://www.startupjournal.com/columnists/enterprise/20050224-bounds.html>
- [Bower, 1981] Bower, G. H., and Hilgard, E. R. (1981). *Theories of learning* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- [Bradshaw, 1999] Bradshaw, J. M., Greaves, M., Holmback, H., Jansen, W., Karygiannis, T., Silverman, B., Suri, N., and Wong, A. (1999). Agents for the masses: Is it possible to make development of sophisticated agents simple enough to be practical? *IEEE Intelligent Systems* (v. 14:2, March 1999), pp. 53-63.
- [Brase, 2003] Brase, J., and Nejd, W. (20003). *Ontologies and Metadata for eLearning*, pages 579–598. Springer Verlag, 2003.
- [Brasher, 2003] Brasher, A. and McAndrew, P. (2003) Metadata vocabularies for describing learning objects: implementation and exploitation issues. *Learning Technology*, 2003. 5(1).
- [Brin, 1998] Brin, S. and Page, L., (1998). The PageRank Citation Ranking: Bringing Order to the Web. *Proceedings of ASIS'98*, 161-172.
- [Brown, 1989] Brown, J. S., Collins, A., and Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- [Bruckman, 1999] Bruckman, A. (1999) The day after Net Day: Approaches to educational use of the Net. *Convergence* 5, 1, 24-26.
- [Brusilovsky, 1998] Brusilovsky, P., Eklund, J. and Schwarz, E. (1998), Web-based education for all: A tool for developing adaptive courseware. *Computer Networks and ISDN Systems* 30(1-7), 291-300.
- [Brusilovsky, 1999] Brusilovsky, P. (1999). Adaptive and Intelligent Technologies for Web-based Education. In Rollinger, C.; Peylo, C. (eds.), *Special Issue on Intelligent Systems and Teleteaching, Künstliche Intelligenz*, 4, 19-25.
- [Brusilovsky, 2001] Brusilovsky, P. (2001). Adaptive hypermedia. *User Modeling and User-Adapted Interaction*, 11, 87–110.
- [Brusilovsky, 2002] Brusilovsky, P. and Maybury, M. (2002). *The Adaptive Web*. *Communications of the ACM*, 2002.

- [Budzik, 2000] Budzik, J. and Hammond, K.J. (2000) User interactions with everyday applications as context for just-in-time information access. In Proceedings of the 2000 International Conference on Intelligent User Interfaces, New Orleans, Louisiana, 2000. ACM Press.
- [Calvanese, 1998] Calvanese, D., De Giacomo, G., and Lenzerini, M. (1998). On the decidability of query containment under constraints. In Proc. of the 17th ACM SIGACT SIGMOD SIGART Symp. on Principles of Database Systems PODS'98), pages 149-158, 1998.
- [Carliner, 2000] Carliner, S. (2000). Physical, cognitive, and affective: a three-part framework for information design. Technical communication. 47(4), 561-576.
- [Carr, 2000] Carr, M., McGee, C., Jones, A., McKinley, E., Bell, B., Barr, H. and Simpson, T. (2000). Strategic research: initiative literature review: the effects of curricula and assessment on pedagogical approaches and on Educational Outcomes. Wellington, Ministry of Education.
- [Chambers 1999] Chambers J. (2000). Keynote talk to the Fall 1999 Comdex Trade Show, Las Vegas, November 16th, 1999.
- [Chang, 2002] Chang, K., Chen, I., and Sung, Y. (2002). The effect of concept mapping to enhance text comprehension and summarization. The Journal of Experimental Education 71(1), 5-23.
- [Chee, 2004] Chee, Y.S. (2004). Distance Education and E-learning in the Digital Age: Critical Considerations. In T. Shih (ed.) (2004), Intelligent Virtual World: Technologies and Applications in Distributed World Environments, pp.289-308. World Scientific Publishers.
- [Chen, 2005] Chen, C. M., Lee, H. M., Chen, Y. H., (2005). Personalized e-learning system using item response theory. Computers and Education (44):237-255.
- [Cheung, 2003] Cheung, K.W., Kwok, J.T., Law, M.H. and Tsui, K.C. (2003) Mining customer product ratings for personalized marketing. Decision Support Systems Vol. 35 (2003), pp. 231-243.
- [Chidlovskii, 2000] Chidlovskii, B., Glance, N. and Grasso A. (2000). Collaborative Re-Ranking of Search Results. Proceedings of AAAI-2000 Workshop on Artificial Intelligence for Web Search, Austin, Texas, 18-22.
- [Cho, 2006] Cho, K., Chung, R., King, W., and Schunn, C. D. (2006). Creating Peer-Based Knowledge Refineries. Proceedings of 2006 Academy of Management Meeting, August 11-16, Atlanta, Georgia.

- [Chou, 2000] Chou, C., Chang, Y. F. and Jiang, Y. Y. (2000). The Development of an Online Adaptive Questionnaire for Health Education in Taiwan. *Computers and Education*, 35, 209-222.
- [Chung, 2005] Chung, R., Bateman, P., and Cho, K. (2005). Expertise gaps and profiles: An integrated view of expertise on knowledge transfer. *Proceedings of IS-CORE 2005 Workshop of International Conference of Information Systems*, Dec 11-14 2005. Las Vegas.
- [Cicognani, 2000] Cicognani, A. (2000). Concept Mapping as a Collaborative Tool for Enhanced Online Learning, *Journal of Educational Technology and Society and IEEE Learning Technology Task Force*, vol. 3(3), 150-158.
- [Cohn, 1996] Cohn, D. A., Ghahramani, Z., and Jordan, M. I. (1996). Active learning with statistical models. *Journal of Artificial Intelligence Research*, 4, 129.
- [Colkin, 2001] Colkin, E. (2001), "Personalization tools dig deeper", *Information Week*, No. 852, pp. 49-51.
- [Collier, 2002] Collier, G. (2002) "e-Learning application infrastructure" *Eduworks- Sun Microsystems White Paper*
- [Collins, 1993] G. Collins, L. Birnbaum, B. Krulwich, and M. Freed. The Role of Self-Models in Learning to Plan. In A. Meyrowitz and S. Chipman, editors, *Foundations of Knowledge Acquisition: Machine Learning*, pages 117–143. Kluwer Press, 1993.
- [Conlan, 2002a] Conlan, O., Dagger, D., Wade, V. (2002) Towards a Standards-based Approach to e-Learning Personalization using Reusable Learning Objects. *E-Learn 2002, World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education*, Montreal, September 2002.
- [Conlan, 2002b] Conlan, O., Hockemeyer, C., Wade, V., and Albert, D. (2002). Metadata Driven Approaches to Facilitate Adaptivity in Personalized eLearning systems. *The Journal of Information and Systems in Education*, 1, 38–44.
- [Conlon, 2007] Conlan, O., O’Keeffe, I., Brady, A., and Wade, V. (2007). Principles for Designing Activity-based Personalized eLearning. *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)* pp. 642 – 644.
- [Conner, 2003] Conner, M.L. "Andragogy and Pedagogy." *Ageless Learner*, 1997 – 2003. <http://agelesslearner.com/intros/andragogy.html>
- [Cooper, 1999] Cooper, L. (1999). Anatomy of an on-line course. *T.H.E. Journal*, 26, 49. Retrieved May 25, 1999 from *INFOTRAC SearchBank* (Article # A53929573).

[Corporate University Exchange, 2000] Corporate University Exchange 2000, <http://www.corpu.com/>

[Cox, 1994] Cox, M. T., and Ram, A. (1994). Choosing learning strategies to achieve learning goals. In M. desJardins and A. Ram (Eds.), Proc. of the AAAI Spring Symp. on Goal-Driven Learning (pp.12-21), Menlo Park, CA: AAAI Press.

[Cristea, 2003] Cristea, A. (2003). Adaptive Patterns in Authoring Educational Adaptive Hypermedia. *Educational Technology and Society*, 6(4), 1-5. Available at http://ifets.ieee.org/periodical/6_4/1.pdf

[Cuban, 2001] Cuban, L. (2001) *Oversold and Underused: Computers in the Classroom* (Cambridge, Harvard University Press).

[Cunningham, 1999] Cunningham, H. (1999) A Definition and Short History of Language Engineering. *Journal of Natural Language Engineering*, pages 1--16, vol 5, 1999.

[Currier, 2004] Currier, S., et al., Quality Assurance for Digital Learning Object Repositories: Issues for the Metadata Creation Process. *ALT-J, research in Learning Technology*, 2004. 12(1): p. 6-20.

[Dardenne, 1993] Dardenne, A., van Lamsweerde A. and Fickas, S. (1993). Goal-Directed Requirements Acquisition, *Science of Computer Programming*, 20, pp. 3-50, 1993.

[Davis, 1989] Davis, F. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.

[Deerwester, 1990] Deerwester, S., Dumais, S. T., Landauer, T. K., Furnas, G. W. and Harshman, R. A. (1990). Indexing by latent semantic analysis. *Journal of the Society for Information Science*, 41(6), 391-407.

[Dennett, 1987] Dennett, D. C. (1987). *The Intentional Stance*. The MIT Press, 1987.

[DesJardins, 1992] DesJardins, M. (1992) Goal-Directed Learning: A Decision-Theoretic Model for Deciding What to Learn Next. In *Proceedings of the Machine Discovery Workshop of the Ninth International Machine Learning Conference*, pages 147–151, 1992.

[De Bra, 1998] De Bra, P. and Calvi, L. (1998). AHA! An open Adaptive Hypermedia Architecture. *The New Review of Hypermedia and Multimedia* 4, 115-139.

[De Vries, 2005] De Vries, P. (2005), *An analysis framework approach for managing corporate E-learning development*, Universiteit Twente, 2005.

- [Dochy, 1992] Dochy, F. J. R. C. (1992). Assessment of prior knowledge as a determinant for future learning. Doctoral Dissertation. Heerlen: Open University.
- [Driscoll, 2002] Driscoll, M. (2002). Web-Based Training: Designing E-Learning Experiences. Jossey-Bass.
- [Drozdova, 2007] Drozdova, M., and Dado, M. (2007), Innovation in engineering education based on the implementation of e-education. European Journal of Engineering Education, 2007 Volume 32, Issue 2 May 2007, pp. 193 – 201.
- [Drucker, 2000] Drucker P. (2000). Need to Know: Integrating e-Learning with High Velocity Value Chains, A Delphi Group White Paper, <http://www.delphigroup.com/pubs/whitepapers/20001213-e-learning-wp.pdf>.
- [Dunlap, 2003] Dunlap, J. C., and Grabinger, R. S. (2003). Preparing students for lifelong learning: A review of instructional features and teaching methodologies. Performance Improvement Quarterly, Volume 16 / Number 2.
- [Ehlers, 2004] Ehlers, U-D. (2004). Quality in e-learning from a learner's perspective. European Journal of Open and Distance Learning, 2004-I, 2004.
- [Ellis, n. d.] Ellis, E., Larkin, M ., and Worthington, L. Executive summary of the research synthesis on effective teaching principles and the design of quality tools for educators. University of Alabama, AL.
- [Eric, 2000] Eric J. Glover, Steve Lawrence, Michael D. Gordon, William P. Birmingham, and C. Lee Giles (2000). Web search – your way. Communications of the ACM, 2000.
- [Farrand, 2002] Farrand, P., Hussain, F. and Hennessy, E. (2002). The efficacy of the 'concept map' study technique. Medical Education 36:426-431.
- [Felder, 1988] Felder, R. M. (1988). Learning and Teaching Styles in Engineering Education. Engr. Education, 78(7), 674–681.
- [Felder, 1993] Felder, R. M. (1993). Reaching the Second Tier: Learning and Teaching Styles in College Science Education. J. College Science Teaching, 23(5), 286-290.
- [Ferragina, 2005] Ferragina, P. and Gulli, A. (2005). A Personalized Search Engine Based on Web-Snippet Hierarchical Clustering. In the Proceedings of the 14th international conference on World Wide Web, Chiba, Japan, 801-810.
- [Firdiyek, 1999] Firdiyek, Y. (1999). Web-based courseware tools: where is the pedagogy? Educational Technology, 39(1), 29-34.

-
- [Flavell 1976] Flavell J. H. (1976). "Metacognitive aspects of problem solving", in L. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.
- [Fleer, 1992] Fleer, M. (1992). Identifying teacher-child interaction which scaffolds scientific thinking in young children. *Science Education*, 76(4), 373-397.
- [Flick, 2000] Flick, L. B. (2000). Cognitive scaffolding that fosters scientific inquiry in middle level science. *Journal of Science Teacher Education*, 11(2), 109-219.
- [Fritch, 2003] Fritch J. W. (2003). Heuristics, tools, and systems for evaluating Internet information. *Online Information Review*. Volume 27 Number 5, 2003, pp. 321-327.
- [Fu, 2000] Fu, X., Budzik, J. and Hammond, J. K. (2000). Mining Navigation History for Recommendation. *Proceedings of the 2000 International Conference on Intelligent User Interfaces*, 106-112.
- [Gagne, 1992] Gagne, R., Briggs, L., Wager, W. (1992). *Principles of Instructional Design*. New York: Holt, Rinehart and Winston.
- [Garrison, 1987] Garrison, D.R., and Shale, D. (1987). Mapping the boundaries of distance education: Problems in defining the field. *The American Journal of Distance Education*, 1(1), 7-13.
- [Garrison, 1997] Garrison, D.R. (1997). Self-directed learning: Toward a comprehensive model. In *Adult Education Quarterly*, Fall 97 v 48 n 1, p18, 16 p.
- [Garro, 2002] Garro (2002), An XML Multi-Agent System for e-Learning and Skill Management, *Third International Symposium on Multi-Agent Systems, Large Complex Systems, and E-Businesses (MALCEB'2002)*, Erfurt, Thuringia (Germany).
- [Glaser, 1962] Glaser, R. (1962). *Psychology and Instructional Technology*. In R. Glaser (Ed.), *Training Research and Education*. Pittsburg: University of Pittsburg Press.
- [Good, 1990] Good, T. L. and Brophy, J. E. (1990). *Educational psychology: A realistic approach*. (4th ed.). White Plains, NY: Longman.
- [Gotschall, 2000] Gotschall, M.. (2000). E-learning strategies for executive education and corporate training. *Fortune*, 141 (10) S5 – S59.
- [Graham, 2007] Graham, A. (2007). The Personal Learning Environments - the future of eLearning? *eLearning Papers*, vol. 2 no. 1.
- [Greenagel, 2002] Greenagel, F. L. (2002). The Illusion of e-learning: Why We Are Missing Out on the Promise of Technology. *League for Innovation in the Community College White paper*.
-

[Guri-Rosenblit, 2005] Guri-Rosenblit, S. (2005). Eight Paradoxes in the Implementation Process of E-learning in Higher Education. *Higher Education Policy*, 2005, 18, (5–29).

[Gustafson, 1991] Gustafson, K. L. (1991). Survey of instructional development models. Syracuse, NY: ERIC Clearinghouse on Information Resources.

[Hall, 1997] Hall, B. (1997). Web-based training cookbook. New York: Wiley.

[Hammond, 1993] Hammond, K., Converse, T., Marks, M., and Seifert, C. M. (1993). Opportunism and Learning. *Machine Learning*, 10(3):279–309, 1993.

[Hanna, 1998] Hanna, D.E. (1998) Higher education in an era of digital competition: Emerging organizational methods. *Journal of Asynchronous Learning Networks*, volume 2, number 1.

[Harasim, 1995] Harasim, L., Hiltz, S. R., Teles, L. and Turoff, M. (1995). *Learning Networks: A Field Guide to Teaching and Learning Online*. Cambridge, MA: The MIT Press.

[Harman, 2007] Harman, K. and Koohang, A. (Eds.) (2007). *Learning Objects, Applications, Implications and Future Direction*. Santa Rosa, California: Informing Science Press. (pp. 1-34)

[Hartman, 2002] Hartman, H. (2002). Scaffolding and Cooperative Learning. *Human Learning and Instruction* (pp. 23-69). New York: City College of City University of New York.

[Haase, 2004] Haase, K. (2004). Context for Semantic Metadata Proceedings. *ACM Multimedia 2004*, New York, 2004.

[Hennessy, 2005] Hennessy, S., Ruthven, K. and Brindley, S. (2005). Teacher perspectives on integrating ICT into subject teaching: commitment, constraints, caution, and change. *Journal of Curriculum Studies*, vol. 37 (2), 155-192.

[Hentea, 2003] Hentea, M., Shea, M., Pennington, L. (2003). A Perspective on Fulfilling the Expectations of Distance Education, In the proceeding of the 4th conference on information technology curriculum on Information technology education, Lafayette, Indiana, USA, 2003.

[Herlocker, 2001] Herlocker, J. and Konstan, J. (2001). Content-Independent, Task-Focused Recommendations. *IEEE Internet Computing*, 5, 40-47.

[Hodgins, 2002] Hodgins, H. W. (2002). The future of learning objects. In D. A. Wiley (Ed.), *The Instructional Use of Learning Objects* (pp. 281-298). Bloomington, Indiana: AIT/AECT.

- [Holman, 1997] Holman, D., Pavlica, K. and Thorpe, R. (1997). Rethinking Kolb's theory of experiential learning in management education, *Management Learning*, Sage, London.
- [Honey, 1986] Honey, P., and Mumford, A. (1986). *Using your Learning Styles*: Peter Honey, Maidenhead.
- [Hong, 2007] Hong, C.M., Chen, C.M., Chang, M.H., and Chen, S.C. (2007). Intelligent Web-based Tutoring System with Personalized Learning Path Guidance. *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)* pp. 512 – 516.
- [Horton, 2000] Horton, W.K. (2000). *Designing Web-Based Training: How to Teach Anyone Anything Anywhere Anytime*. John Wiley and Sons.
- [Huang, 2007] Huang, M.J., Huang, H.S., and Chen, M.Y. (2007). Constructing a personalized e-learning system based on genetic algorithm and case-based reasoning approach. *Expert System Applications*, Volume 33, Issue 3, October 2007, pp. 551 – 564.
- [Huget, 2004] Huget, M.P. and Odell, J. (2004). Representing agent interaction protocols with agent UML. *Proceedings of the Fifth International Workshop on Agent Oriented Software Engineering (AOSE)*.
- [IEEE, n.d.] IEEE, IEEE Standard for Learning Object Metadata, in IEEE Std 1484.12.1-2002. 2002, The Institute of Electrical and Electronics Engineers, Inc. p. i-32.
- [iN2015, 2006] iN2015 Education and Learning Report. Retrieved July 30, 2006, from <http://www.in2015.sg/reports.html>
- [Jansen, 2000] Jansen, B., Spink, A., Saracevic, T. (2000). Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web. *Information Processing and Management* 36(2): 207-227.
- [Jennings, 1998] Jennings, N.R., Sycara, K., and Wooldridge, M. (1998) 'A roadmap of agent research and development', *Autonomous Agents and Multi-Agent Systems*, 1, 7-38.
- [Johnson, 1986] Johnson, R. T., and Johnson, D. W. (1986). Action research: Cooperative learning in the science classroom. *Science and Children*, 24, 31-32.
- [Jonassen, 1991] Jonassen, D. H. (1991), Objectivism vs constructivism: Do we need a new philosophical paradigm. *Educational Technology Research and Development*, 39(3), 5-14.
- [Jonassen, 1992] Jonassen, D.H. (1992). Applications and limitations of hypertext technology for distance learning. Paper presented at the Distance Learning Workshop, Armstrong Laboratory, San Antonio, TX.

[Kasanoff, 20001] Kasanoff, B. (2001), Making it Personal, Perseus Publishing, Cambridge, MA.

[Kao, 2001] Kao, T. T. (2001). Personalization Courseware Navigation Using Grey Ranking Analysis, Master's thesis, Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan.

[Karl, 2006] Karl, M. (2006). Embracing Informal Learning. White paper for Department of Instructional Technology and Institute for Interactive Technologies.

[Karon, 2000] Karon, R. L. (2000). Bank solves compliance training challenge with internet. Elearning, January-March.

[Kearn, 2000] Kearn C. M. (2000). Affecting the Future: The Role of. Appropriate Scaffolding in the. Development of Social Competence. In Dianne Rothenberg, Ed. 2002. Issues in Early Childhood Education: Curriculum, Teacher Education, and Dissemination of Information. Proceedings of the Lilian Katz Symposium, November 5-7, 2000

[Keegan, 1986] Keegan, D. (1986). The foundations of distance education. London: Croom Helm.

[Keller, 2004] Keller, J.M. and Suzuki, K. (2004). Learner motivation and E-learning design: a multinationally validated process. Journal of Educational Media, Vol. 29, No. 3, October 2004.

[Kelly, 2003] Kelly, A.E., (Ed.). (2003). Theme issue: The role of design in educational research. Educational Researcher, 32 (1).

[Kerr, 1991] Kerr, S.T. (1991) Lever and fulcrum: educational technology in teachers' thought and practice, Teachers College Record, 93(1), pp. 114-136.

[Khan, 1993] Khan, K. M. (1993). Concept Mapping as a Strategy for Teaching and Developing the Caribbean Examinations Council (CXC) Mathematics Curriculum in a Secondary School. The Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Ithaca, NY, Misconceptions Trust

[Khan, 1997] Khan, B. H. (1997). Web Based Instruction. Englewood Cliffs, New Jersey: Educational Technology Publications.

[Kleinberg, 1998] Kleinberg, J. (1998). Authoritative Sources in a Hyperlinked Environment. Proceedings of 9th ACM-SIAM Symposium on Discrete Algorithms, 668-679.

[KnowNet, 2004] KnowNet (2004). Pedagogy, e-learning and knowledge development. Open Source for Collaborative Knowledge Development and Learning.

[Ko, 2005] Ko, D., Kirsch, L.J., and King, W.R. (2005). "Antecedents of knowledge transfer from consultants to clients in enterprise system implementations," *MIS Quarterly* (29:1) 2005, pp 59-85.

[Kobayashi, 2000] Kobayashi, M. and Takeda, K. (2000). Information Retrieval on the Web. *ACM Computing Surveys*, 32(2), 144-173.

[Kobsa, 1993] Kobsa, A. (1993). User modeling: Recent Work, Prospects and Hazards. In *Adaptive User Interfaces: Principles and Practice*, M. Schneider-Hufschmidt, T. Kühme, and U. Malinowski, (eds.). 1993, North-Holland: Amsterdam.

[Kobsa, 2002] Kobsa, A. (2002). Personalized hypermedia and international privacy. *Communications of the ACM* 45, 5 (2002), 64–67.

[Kolb, 1984] Kolb, D.A. (1984) *Experiential Learning: experience as the source of learning and development* New Jersey: Prentice-Hall.

[Konstan, 1997] Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L. and Riedl, J. (1997). GroupLens: applying collaborative filtering to usenet news. *Communications of the ACM*, 40(3), 1997, pp. 77-87.

[Koper, 2003] Koper, E. J. R., and Sloep, P. (2003). *Learning Networks: connecting people, organizations, autonomous agents and learning resources to establish the emergence of effective lifelong learning* (OTEC RTD Programme Plan 2003-2008). Heerlen: Open Universiteit Nederland.

[Koper, 2004] Koper, R., and Olivier, B. (2004). Representing the Learning Design of Units of Learning. *Educational Technology and Society* 7 (3), 97-111.

[Koper, 2004a] Koper R. (2004). Use of the Semantic Web to Solve Some Basic Problems in Education, *Journal of Interactive Media in Education*, (1), Special Issue on the Educational Semantic Web.

[Kwok, 2004] Kwok L.Y., Tan Y. G. (2004). Scaffolding Supports in Project-based Learning through Knowledge Community (KC): collaborative learning strategies and pedagogical facilitation. 8th GCCCE2004 Conference.

[Laroussi, n. d.] Laroussi, M. (n. d.). New e-learning services based on mobile and ubiquitous computing UBI-Learn Project, Retrieved November 30, 2004, from http://www-clips.imag.fr/calie04/actes/Laroussi_final.pdf

[Lawrence, 1998] Lawrence, S. and Giles, C. L. (1998). Searching the World Wide Web. *Science*, 280(5360), 98-100. Lawrence, S. and Giles, C. L. (1999). Searching the Web: General and Scientific Information Access. *IEEE Communications*, 37, 116-122.

-
- [Lawrence, 2000] Lawrence, S. (2000). Context in Web Search, IEEE Data Engineering Bulletin, Volume 23, Number 3, pp. 25–32, 2000.
- [Lee, 2001] Lee, M-G. (2001). Profiling Students' Adaptation Styles in Web-based Learning. Computers and Education, 36, 121-132.
- [Lin, 2001] Lin, H. N. (2001). An Adaptive Learning Environment: A Case Study on High School Mathematics, Master's Thesis, Department of Computer and Information Science, National Chiao Tung University, Hsinchu, Taiwan.
- [Lin, 2003] Lin, Y., and Ding, H. (2003). Modeling Probe-Agent Distributed System by Goal-oriented Approach. CAiSE Short Paper Proceedings, 2003.
- [Lin, 2006] Lin, F., Holt, P., Leung, S., and Li, Q. (2006). A multiagent and service-oriented architecture for developing adaptive e-learning systems. International Journal of Continuing Engineering Education and Life Long Learning. Volume 16, Number 1-2 / 2006.
- [Lim, 2001] Lim, G. (2001). A Framework for Integrating Technologies in Teaching and Learning. 12th International Conference for Society of Information Technology and Teacher Education (SITE2001), Florida, USA, 7 March 2001.
- [LOM, 2000] LOM (2000). LOM working draft v4.1 [online]. Available: <http://ltsc.ieee.org/doc/wg12/LOMv4.1.htm>
- [LTSC, 2000] LTSC (2000). IEEE P1484.12 Learning Objects Metadata Working Group homepage.
- [Lu, 2004] Lu, J. (2004). A personalized e-learning material recommender system, Proceedings of ICIT 2004, pp. 374-379.
- [Lynch, 2001] Lynch, O. (2001) Welcome Speech Building an ICT Research Network: Helping to Create Schools of the Future (Barbican Centre, London, Becta).
- [Lytras, 2002] Lytras, M. (2002). E-learning pedagogy: The reveal of value adding learning processes. Definitions and Implications for dynamic learning content delivery. *World Conference on E-Learning in Corp., Govt., Health., and Higher Ed. 2002(1)*, 1826-1830.
- [Lytra, 2001] Lytra, M.D. and A. Pouloudi (2001). E-learning: Just a Waste of Time. In Strong, D., Straub, D. and DeGross, J.I. (Eds) Proceedings of the Seventh Americas Conference on Information Systems (AMCIS), Boston, Massachusetts, pp. 216-222.
- [Lytras, 2001a] Lytras M. D., Doukidis, G. I. and Skagkos, T. N. (2001). E-learning pedagogy: A value definition from a knowledge management perspective. In Technology in Education International Conference and Exposition (TechEd 2001), 2001.
-

[Mager, 1962] Mager, R.F. (1962). Preparing Objectives for Programmed Instruction. Belmont, CA: Fearon.

[Malcolm, 2005] Malcolm M. (2005). The exercise of the object: issues in resource reusability and reuse. Br. J. of Educ. Technology Vol. 36 No. 1, pp. 33-41.

[Malopinsky, 2000] Malopinsky, L., Kirkley, J., Stein, R., and Duffy, T. (2000). An instructional design model for online problem based learning (PBL) environments: The Learning to Teach with Technology Studio. Paper presented at the Association for Educational Communications and Technology Conference (AECT), October 26, Denver, Colorado.

[Maltz, 2001] Maltz, Maxwell (2001) New Psycho-Cybernetics. Prentice Hall. ISBN 0-7352-0275-3

[Manjunath, 2006] Manjunath, B., Patil, S. (2006). E-Learning and Its Impact on Library and Information Services. DRTC Conference on ICT for facilitating Digital Learning Environment, Bangalore, India, 11-13 January, 2006.

[Martinez, 2000] Martinez, M. (2000). Successful Web learning environments: New design guidelines (ERIC Document Reproduction Service No. ED 446 745).

[Mason, 2006] Mason, R. (2006). E-learning in Tertiary Education: Where do we stand? Centre for Educational Research and Innovation, Higher Education Quarterly, Volume 60, Number 3, July 2006 , pp. 287-289(3).

[Maurer, 2001] Maurer, H. and Sapper, M. (2001) E-learning has to be seen as part of general knowledge management. In Proceedings of ED-MEDIA 2001, pages 1249–1253, Charlottesville, USA, 2001. AACE.

[Mayer, 1972] Mayer, R., and Greeno, J. G. (1972). Structural differences between learning outcomes produced by different instructional methods. Journal of Educational Psychology, 63, 165-173.

[McArdle, 2005] McArdle G., Monahan T., Bertolotto M. and Mangina E. (2005). Analysis and Design of Conceptual Agent Models for a Virtual Reality e-Learning Environment, International Journal on Advanced Technology for Learning, ACTA Press, 2005.

[McGee, 2002] McGee, P. and Wickersham, L. (2002). The WWW, ADDIE, and the ADL Guidelines. In G. Richards (Ed.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2002 (pp. 2667-2668). Chesapeake, VA: AACE.

-
- [McGreal, 2001] McGreal, R., and Roberts, T. (2001). A Primer on Metadata for Learning Objects. Availability: <http://www.elearningmag.com/issues/Oct01/learningobjects.asp>
- [McKenzie, 1999] McKenzie, J. (1999). Scaffolding for Success. Beyond Technology, Questioning, Research and the Information Literate School Community.
- [McLeod, n. d.] McLeod, G. (n. d.) Learning Theory and Instructional Design.
- [McNee, 2002] McNee, S.M., Albert, I., Cosley, D., Gopalkrishnan, P., Lam, S.K., Rashid, A.M., Konstan, J.A and Riedl, J. (2002) On the recommending of citations for research papers. In Proceedings of ACM International Conference on Computer Supported Collaborative Work (CSCW'02), 116-125.
- [Merriam, 1991] Merriam, S. and Caffarella (1991, 1998) *Learning in Adulthood. A comprehensive guide*, San Francisco: Jossey-Bass.
- [Merriam, 1999] Merriam, S. B. and Caffarella, R. S. (1999). Learning in Adulthood: A Comprehensive Guide. (2nd Edition). San Francisco: Jossey-Bass.
- [Merrill, 1983] Merrill, M. D. (1983). Component Display Theory. In C. Reigeluth (Ed.), Instructional Design Theories and Models. Hillsdale, NJ: Erlbaum Associates (1983).
- [Merrill, 2001] Merrill, M. David. (2001). Components of Instruction Toward a Theoretical Tool for Instructional Design. Instructional Science, 29, 291-310.
- [Meyer, 1993] Meyer, D. K. (1993). What is scaffolded instruction? Definitions, distinguishing features, and misnomers. In D. J. Leu and C. K. Kinzer, (Eds.), Examining central issues in literacy research, theory, and practice (pp. 41-53). Chicago: The National Reading Conference.
- [Michalski, 1995] Michalski, R. S., and Ram, A. (1995). Learning as goal-driven inference. In A. Ram and D. Leake (Eds.), Goal-driven learning (pp. 479-490). Cambridge, MA: MIT Press/Bradford Books.
- [Miettinen, 2000] Miettinen, R. (2000). The concept of experiential learning and John Dewey's theory of reflective thought and action, International Journal of Lifelong Education, 19 (1), January-February, pp 54-72
- [Millard, 2005] Millard, D., Woukeu, A., Tao, F. B. and Davis, H. (2005) The Potential of Grid for Mobile e-Learning. In Proceedings of The 4th World Conference on Mobile Learning (MLEARN 2005), Cape Town, South Africa.
- [Mobasher, 2000] Mobasher, B., Cooley, R. and Srivastava, J. (2000). Automatic Personalization Based on Web Usage Mining. Communications of the ACM, 43(8), 142-151.
-

- [Mor, 2004] Mor, E. and Minguillo'n, J. (2004), "E-learning Personalization based on Itineraries and long-term navigational behavior", Proceedings of the 13th International World Wide Web Conference, Alternate Track Papers and Posters, pp. 264-265.
- [Morris, 2002] Morris, H. and Rippin, A. (2002) E-learning in Business and Management: The Current State of Play. BEST 2002 Conference, Support the Teacher: Challenging the Learner, Edinburgh.
- [Morrison, 2001] Morrison, G. R., Ross, S. M. and Kemp, J. E. (2001). Designing Effective Instruction. (3rd Edition). New York: John Wiley and Sons.
- [Moursund, 1999] Moursund, D. (1999). Project-based learning using IT. Eugene, Or. : International Society for Technology in Education.
- [Nadolski, 2007] Nadolski, R., Hummel, H., Brink, H van den, and Hoefakker, R. (2007). EMERGO: methodology and toolkit for efficient development of serious games in higher education. Simulation and Gaming.
- [Najberg, 2007] Najberg, B. L. (2007). Building Models in Instructional Design. Instructional Design and Learning Newsletter. Vol. 2, Issue 4.
- [Narayanan, 2007] Narayanan, S., Lim, L. and Tan, M. (2007). Courseware Administration System – QC Made Easy. In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2007 (pp. 1537-1546). Chesapeake, VA: AACE.
- [NUS, 2007] National University of Singapore (2007), Web 2.0 and Personal Learning Environments, IDEAS, Centre for Instructional Technology's newsletter on enhancing teaching and learning with digital and AV technology, April, 2007.
- [Newman, 1993] Newman, F., and Holzman, L. (1993). Lev Vygotsky: Revolutionary scientist. New York: Routledge.
- [Newton, 2001] Newton, L.R. and Rogers, L. (2001) Teaching Science with ICT (London, Continuum).
- [Ngeow, 2001] Ngeow, K.K., and Yoon, S. (2001). Learning to learn: preparing teachers and students for problem-based learning. ERIC Digest.
- [Nielsen, 1998] Nielsen, J. (1998), "Personalization is over-rated", Alertbox. available at: www.useit.com/alertbox/981004.html
- [Novak, 1984] Novak, J.D. and Gowin, D.B. (1984). Learning how to learn. Cambridge: Cambridge University Press.

- [Noy, 2001] Noy, N. and McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*. Knowledge Systems Laboratory, March, 2001.
- [O'Hear, 2006] O'Hear, S. (2006). e-learning 2.0: How Web technologies are shaping education (3 September, 2007) http://www.readwriteweb.com/archives/e-learning_20.php
- [OCLC, 2003] OCLC E-Learning Task Force. (2003). Libraries and the enhancement of e-learning. Retrieved July 21, 2004, from <http://www5.oclc.org/downloads/community/elearning.pdf>
- [Ofsted, 1999] Ofsted (1999) *Standards in the Secondary Curriculum 1997/98: Information Technology*(London, Office for Standards in Education).
- [Okamoto, 2005] Okamoto, K. (2005) *Creating New e-Pedagogy between Personalized e-Learning and Computer Supported Collaborative Learning*.
- [Palincsar, 1998] Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, 49, 345-375.
- [Pankratius, 2003] Pankratius, V., and Vossen, G. (2003) *Towards E- Learning Grid:Using Grid Computing in Electronic Learning*.Canada:Proc IEEE workshop on Knowledge Grid and Grid Intelligence, 2003.
- [Papanikolaou, 2002] Papanikolaou K. A. and Grigoriadou, M. (2002). Towards New Forms of Knowledge Communication: the Adaptive Dimensions of a Web-based Learning Environment. *Computers and Education*, 39, 333-360.
- [Park, 2000] Park, K., Kim, J., and Park, S. (2000). Goal based agent-oriented software modeling. *apsec*, p. 320, Seventh Asia-Pacific Software Engineering Conference (APSEC'00), 2000.
- [Parker, 2004] Parker, N. K., (2004), "The Quality Dilemma in Online Education" in *Theory and Practice of Online Learning*, Athabaska, Athabaska University Press
- [PBS, 2001] PBS (2001). PBS Distance Learning Week, April 15 - 21. PBS
- [Perraton, 1988] Perraton, H. (1988). A theory for distance education. In D. Sewart, D. Keegan, and B. Holmberg (Ed.), *Distance education: International perspectives* (pp. 34-45). New York: Routledge.
- [Phillips, 2005] Phillips, R. A. (2005). Challenging the primacy of lectures: The dissonance between theory and practice in university teaching. *Journal of University Teaching and Learning Practice*, 2(1), 1.

[Pitkow, 2002] Pitkow, J., Schutze, H., Cass, T., Cooley, R., Turnbull, D., Edmonds, A., Adar, E., and Breuel, T. (2002). Personalized Search. *Communications of the ACM*, Vol. 45(9):pages 50–55, September 2002.

[Plomp, 1993] Plomp, T., Feteris, A., Pieters, J. M. , and Tomic, W. (Eds.) (1992). *Designing education and training [Ontwerpen van onderwijs en trainingen]*. Utrecht, The Netherlands: Lemma.

[Quitana, 2004] Quitana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D. C., and Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, 13(3), 337-386.

[Rashid, 2002] Rashid, A.M., Albert, I., Cosley, D., Lam, S.K., McNee, S., Konstan, J.A., and Riedl, J. (2002). Getting to Know You: Learning New User Preferences in Recommender Systems. *Proceedings of the 2002 International Conference on Intelligent User Interfaces*, 127-134.

[Rasmussen, 1998] Rasmussen, K. L. (1998). Hypermedia and learning styles: Can performance be influenced? *Journal of Multimedia and Hypermedia*, 7(4).

[Reigeluth, 1983] Reigeluth, C. M., Stein, F. S. (1983). The Elaboration Theory of Instruction. In C. M. Reigeluth (Ed), *Instructional Design Theories and Models: An Overview of their Current States*. Hillsdale, NJ: Lawrence Erlbaum (1983).

[Reigeluth, 1996] Reigeluth, C. M. (1996). A new paradigm of ISD? *Educational Technology and Society*, 36(3), 13-20.

[Reiser, 1987] Reiser, R.A. (1987). Instructional technology: A history. In R. M. Gagn (Ed.), *Instructional technology: Foundations* (pp. 11-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

[Renger, 2002] Renger, W.J., Bijk E. and A. Thomassen MA. (2002). e-Learning and Arts Pedagogy. 7th ELIA Conference, "Comhar", Dublin, Ireland 24-26 October 2002.

[Reynolds, 1997] Reynolds, M (1997) Learning styles: a critique, *Management Learning*, 28 (2) pp 115-33, Sage, London.

[Riding, 1998] Riding, R., and Rayner, S. (1998). *Cognitive Styles and learning strategies*. London: David Fulton Publishers.

[Riecken, 2000] Riecken, D. (2000), "Personalized views of personalization", *Communications of the ACM*, Vol. 43, No. 8, pp. 27-8.

[Robbings, 2002] Robbings, S.R. (2002). The Evolution of the Learning Content Management System. *ASTD's sources for e-learning*.

- [Rogers, 1996] Rogers, A. (1996). *Teaching Adults* (2nd ed.). Buckingham: Open University Press.
- [Rosen, 2006] Rosen, A. (2006). Technology Trends: e-learning 2.0. The elearning Guild's Learning Solutions e-Magazine, October, 2006.
- [Rosenberg, 2001] Rosenberg, M.J. (2001). *E-learning: Strategies for Delivering Knowledge in the Digital Age*. NY: McGraw-Hill.
- [Rowan, 2001] Rowan, B., Schilling, S. G., Ball, D. L. and Miller, R. (2001), *Measuring Teachers' Pedagogical Content Knowledge in Surveys: An Exploratory Study*. Consortium for Policy Research in Education, Study of Instructional Improvement, Research Note S-2. Ann Arbor: University of Michigan.
- [Rowntree, 1992] Rowntree, D. (1992) *Exploring open and distance learning* (London, Kogan Page).
- [Rumetshofer, 2003] Rumetshofer, H., and Wöß, W. (2003). XML-based Adaptation Framework for Psychological-driven E-learning Systems. *Educational Technology and Society*, 6 (4), 18-29, Available at http://ifets.ieee.org/periodical/6_3/4.pdf
- [Saber, 2005] Saber, J., Flores, I., Fagan, M. H., Kilmon, C., Williams, J., Ibitayo., K. (2005). If you build it, will they come? Challenges in E-learning Delivery System Choice. *Issues in Information Systems*. Vol. VI, No. 1, 2005.
- [Schafer, 2001] Schafer, J.B., Konstan, J.A. and Riedl, J. (2001) "E-commerce recommendation applications" *Data Mining and KnowledgeDiscovery*, Vol. 5 (2001), pp. 115–153.
- [Schafer, 1999] Schafer J. B., Konstan, J., and Riedl, J. (1999). Recommender systems in e-commerce. In *Proceedings of the ACM Conference on Electronic Commerce (EC-99)*, pages 158–166, November 1999.
- [Schank, 2001] Schank, R.C. (2001). *Designing World-Class E-Learning*. McGraw-Hill Professional Publishing.
- [Scienter, 2001] Scienter, Bologna 2001.
- [Schreiber, 1998] Schreiber, D. A., and Berge, Z. L. (1998). *Distance training: How innovative organizations are using technology to maximize learning and meet business objectives*. San Francisco: Jossey-Bass.
- [Schunk, 1991] Schunk, D. H. (1991). *Learning theories: An educational perspective*. New York: Macmillan

- [Shen, 2003] Shen, Z.Q. (2003). Goal-Based Intelligent Agent and its Applications, doctorate dissertation, Nanyang Technological University, Singapore.
- [Shen, 2003a] Shen, Z.Q., Gay, R. and Tao, X.H. (2003) "Goal-Based Intelligent Agents", International Journal of Information Technology, Vol 9, No. 1, pp. 19-30, 2003.
- [Shen, 2005] Shen, Z.Q., Gay, R., and Miao, C.Y., Tan, T.W., Kuay, C.S. and Lee, H.M. (2005), Process Modeling and Automated Execution for Bio-Manufacturing, International Journal of Information Technology, 2005.
- [Shen, 2006] Shen, Z.Q., Miao, C.Y., and Gay, R. (2006), "Goal-oriented Methodology for Agent-oriented Software Engineering," IEICE Transaction on Information and Systems, 2006.
- [Shen, 2007] Zhiqi Shen, Z.Q., Miao, C.Y., Gay, R. and Low, C.P. (2007), "Personalized e-Learning – a Goal Oriented Approach". Proc. 7th WSEAS Int. Conf. on Distance Learning and Web Engineering (DIWEB '07), Beijing, China, Sept. 15th-17th, 2007. pp 304-309.
- [Siemens, 2002] Siemens, G. (2002). Instructional Design in Elearning. Elearnspace, September, 2002.
- [Simon, 1974] Simon, H.A., and Lea, G. (1974). "Problem solving and rule induction: A unified view.," in: *Knowledge and cognition*, Erlbaum, Hillsdale, NJ, 1974, pp. 105-127.
- [Simons, 1990] Simons, P. R. J. (1990). Transfervermogen. [Transfer-ability] Inaugural lecture. Nijmegen: Quick Print.
- [Simons, 1999] Simons, P.R.J. (1999). Transfer of learning: paradoxes for learners. International Journal of Educational Research, 31, 577-589.
- [Snelbecker, 1983] Snelbecker, G. E. (1983). *Learning theory, instructional theory, and psychoeducational design*. New York: McGraw-Hill.
- [Spector, 2000] Spector, J. M. (2000). Toward a Philosophy of Instructional Design. Annual Meeting of the American Educational Research Association. New Orleans, LA: AERA 2000.
- [Stojanovic, 2001] Stojanovic, L., Staab, S., and Studer, R. (2001). E-learning based on the semantic web. In WebNet2001 - World Conference on the WWW and Internet, Orlando, Florida, USA, 2001.
- [Stone, 1998] Stone, C.A. (1998). The metaphor of scaffolding: Its utility for the field of learning disabilities. Journal of Learning Disabilities, 31(4), 344-364.

-
- [Sweller, 1998] Sweller, J. (1998) Cognitive Load During Problem Solving: Effects on learning, *Cognitive Science*, 12 (2) (1988) 257-285.
- [Súilleabhain, 2004] Súilleabhain G. Ó., and Coughlan R. (2004). Student support services for e-learning: collaborative tutoring and authentic assessment. Third EDEN Research Workshop, Carl von Ossietzky University of Oldenburg, Germany, pp. 83-86.
- [Tabak, 2004] Tabak, I. (2004). Synergy: A complement to emerging patterns of distributed scaffolding. *The Journal of the Learning Sciences*, 13(3), 305-335.
- [Tang, 2003] Tang, Y. T. and McCalla, G. (2003) Towards pedagogy-oriented paper recommendation and adaptive annotations for a web-based learning system. Workshop on Knowledge Representation and Automated Reasoning for E-Learning Systems, 18th International Joint Conference on Artificial Intelligence (IJCAI 2003), Acapulco, Mexico. August 9-15, 2003.
- [Tang, 2005] Tang, B. (2005). Interactive e-learning activities to engage learners - A simple classification. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005 (pp. 4092-4097)*. Chesapeake, VA: AACE.
- [Taylor, 2005] Taylor, J., Rodden, T., Tom, Anderson, A., Sharples, M., Luckin, R., Conole, G., and Siraj-Blatchford, J. (2005). An e-learning research agenda. <http://www.epsrc.ac.uk/CMSWeb/Downloads/Other/E-learningResearchAgenda.pdf>
- [Tim, 2000] Tim L. Wentling, Consuelo W., Danielle S., Jennie F., Jason La Fleur, Alaina K. (2000). *The Future of e-Learning: A Corporate and an Academic Perspective*. Knowledge and Learning Systems Group University Of Illinois At Urbana-Champaign, September 2000.
- [Totten, 1991] Totten, S., Sills, T., Digby, A., and Russ, P. (1991). *Cooperative learning: A guide to research*. New York: Garland.
- [Trifonova, 2003] Trifonova A., Ronchetti M. (2003), Where is mobile learning going? E-Learn 2003 Conference, USA.
- [Trifonova, 2004] Trifonova, A., Ronchetti, M. (2004). A General Architecture to Support Mobility in Learning. *Proc. of the 4th IEEE International Conference on Advanced Learning Technologies (ICALT 2004)*, August 30 - September 1, 2004, Joensuu, Finland.
- [Tsai, 2002] Tsai, S., Machado, P. (2002) "E-learning, Online Learning, Web-based Learning, or Distance Learning: Unveiling the Ambiguity in Current Terminology" in *eLearn Magazine*.
- [Tynjälä, 2005] Tynjälä, P. and Häkkinen, P. (2005). E-learning at work: theoretical underpinnings and pedagogical challenges *The Journal of Workplace Learning*, 2005.
-

[Tzeng, 2007] Tzeng, G.H., Chiang, C.H. and Li, C.W. (2007). Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL. *Expert Systems with Applications*. Vol. 32, Issue 4, May 2007, pp. 1028 – 1044.

[Urdan, 2000] Urdan, T. A., and Weggen C. C. (2000). *Corporate e-learning: Exploring a new frontier*. WR Hambrecht + Co.

[Van Der Stuyf, 2002] Van Der Stuyf, R. (2002). Scaffolding as a teaching strategy. *Adolescent Learning and Development*. Section 0500A.

[Vasilyeva, 2007] Vasilyeva, E., Pechenizkiy, M., Gavrilova, T., and Puuronen, S. (2007) Personalization of Immediate Feedback to Learning Styles. *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)* pp. 622 – 624.

[Venkatesh, 2003] Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003) User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.

[Visser, 1998] Visser, L. (1998) The development of motivational communication in distance education support. Report for the Educational Technology Department (Twente, University of Twente, The Netherlands).

[Vygotsky, 1978] Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press. (Original work published in 1930, 1933, 1935).

[Wade, 2002] Wade, V., Carroll, V., Huntley-Moore, S., O'Rourke, B., Schwienhorst, K., and Tangney, B. (2002). *e-Learning for Trinity College: Sub Group Report*.

[Wankat, 1993] Wankat, P.C., and Oreovicz, F.S. (1993). *"Teaching Engineering"*, McGrawHill Inc., New York, 1993.

[Watkins, 2005] Watkins, R. (2005). Preparing learners for online learning. Retrieved January 09, 2006: <http://www.learningcircuits.org/2005/sep2005/watkins.html>

[Watkins, 1999] Watkins, C. and Mortimore, P. (1999), *Pedagogy: what do we know?* in Mortimore, P. (Ed.) *Understanding Pedagogy and its Impact on Learning*. London: Chapman.

[WebCT] The WebCT E-Learning Hub, <http://www.webct.com>

[Webster, 1997] Webster, J. and Hackley, P. (1997). Teaching effectiveness in technology-mediated distance learning. *Academy of Management Journal*, 40(6), 1282-1309.

[Wentling, 2000] Wentling, T., Waight, C., Gallager, J., La Fleur, J., Wang, C., and Kanfer, A. (2000) E-Learning – A review of Literature. University of Illinois, Urbana-Champaign, 2000.

[Wertsch, 1985] Wertsch, J. V., and Stone, C. A. (1985). The concept of internalization in Vygotsky's account of the genesis of higher mental functions. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives* (pp. 162-179). Cambridge, England: Cambridge University Press.

[Wild, 2002] Wild, R.H., Griggs, K.A., Downing, T. (2002), "A framework for e-learning as a tool for knowledge management", *Industrial Management and Data Systems*, Vol. 102 No.7, pp.371-80.

[Wiley, 2000] Wiley, D. (2000). *The Instructional Use of Learning Objects*. Bloomington, IN: AECT.

[Wilson, 2005] Wilson B.G. (2005). Theory and method as tools: reflections on research on the pedagogical uses of ICT in education. *Computers in Human Behavior* 21 (2005) pp. 541–546.

[Winn, 1990] Winn, W. (1990). Some implications of cognitive theory for instructional design. *Instructional Science*, 19, 53-69.

[Winne, 1985] Winne, P. H. (1985). Cognitive processing in the classroom. In T. Husen and T. N. Postlethwaite (Eds.), *The International Encyclopedia of Education* (Vol. 2, pp. 795-808). Oxford: Pergamon.

[Wolpers, 2004] Wolpers, M. (2004). Promoting E-Learning Research and Application Scenarios in Europe, Proc. ICL Conf. 2004, Villach, Austria, September 2004

[Wood, 1976] Wood, D., Bruner, J.S., and Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17, 89-100.

[Woodruff, 2000] Woodruff, A., Gossweiler, R., Pitkow, J., Chi, E. and Card, S.K. (2000) Enhancing a digital book with a reading recommender. In Proc. ACM CHI 2000.153-160. 2000.

[Wooldridge, 1995] Wooldridge, M. and Jennings, N. R. (1995). "Intelligent agents: Theory and practice," *The Knowledge Engineering Review*, vol. 10, no. 2, pp. 115-- 152, 1995.

[Woolfolk, 2004] Woolfolk, A. (2004). *Educational psychology* (9th ed.). Boston, MA: Allyn and Bacon.

- [Wu, 2001] Wu, H., De Kort, E., De Bra, P. (2001). "Design Issues for General Purpose Adaptive Hypermedia Systems". Proceedings of the 12th ACM Conference on Hypertext and Hypermedia, Aarhus, Denmark, 2001.
- [Yan, 1999] Yan, T.W. and Garcia-Molina, H. (1999). The SIFT information dissemination system. ACM Transactions on Database Systems, Vol. 24, No. 4, 1999, pp.529-565.
- [Yu, 1998] Yu, E. and Mylopoulos, J. (1998). Why Goal-Oriented Requirements Engineering. In Proceedings of the 4th International Workshop on Requirements Engineering: Foundations of Software Quality. June 1998, Pisa, Italy. E. Dubois, A.L. Opdahl, K. Pohl, eds. Presses Universitaires de Namur, 1998. pp. 15-22.
- [Zahm, 2000] Zahm, S. (2000). No question about it – e-learning is here to stay: A quick history of the e-learning evolution. E-learning, 1 (1) 44-47.
- [Zaine, 2002] Zaine, O. R. (2002). Building a Recommender Agent for e-Learning Systems. Proc. of Int. Conf. on Computers in Education (ICCE'02).
- [Zang, 2003] Zang, D., and Nunamaker, J. F. (2003). Powering e-learning in the new millennium: An overview of e-learning and enabling technology. Information Systems Frontiers, 5 (2), 207.
- [Zvacek, 1991] Zvacek, S. M. (1991) Effective affective design for distance education, TechTrends, 36(1), 40–43.

APPENDIX A

ADDIE MODEL – ILLUSTRATIONS OF PEDAGOGICAL APPROACH

A1 Illustration of Workflow

Background:

This ADDIE model illustration is based on the Singapore Educational System and follows the Ministry of Education (MOE) learning guidelines (see <http://www.moe.gov.sg/>). Only the first two phases of the ADDIE Model, namely the analysis and the design phases, will be illustrated as all the pedagogical considerations are covered by these two phases. The next three phases, namely development, implementation and evaluation phases, are typically technical in nature as they look at the provision of content from an implementation and delivery point of view. Hence, these phases will not be illustrated.

Requirement:

Develop a science courseware for secondary two students

Topic Chosen: Chapter 9: Introducing Electricity

Chapter 10: Household Electricity

Analysis Phase:

1. Feasibility Analysis

- Conduct Feasibility Report
- Analyze Feasibility Report

The feasibility analysis sets out to justify the use of e-learning as an effective teaching means. Essentially, it helps to sort out all the main issues covering the e-learning solution (Rationale, Stakeholders/Collaborators Involvement and Support, Content, Learner, Pedagogy, Technical Constraints) and determines how practical/feasible it is for the organization to implement an e-learning solution.

It is important to note that the feasibility report is different from a design report in terms of the emphasis placed on the practical and economic aspects of the solution. The main focus of a design report is to describe a specific implementation. A feasibility study,

however, goes a step further. Not only does it describe a specific implementation, it also investigates and compares alternative teaching solutions.

A feasibility report is essential as it serves to “document an engineer's thinking through a solution to a problem, a description of the solution, and the reasons why that solution should be implemented” [Perelman, 1998, p. 80]. Furthermore, it also acts as a springboard for developing other documents, such as proposals, specifications, and work plans. In summary, the key rationales (of the report) to clarify are:

- a. Does the courseware fit with the institutional Learning and Teaching Strategy adopted?
- b. Does the courseware fit with the e-learning strategy of the department (if you have one) or the institution
- c. How can the use of educational technologies be embedded into the student learning experience?
- d. Does a clear strategy exist for how the courseware is to be developed or resourced?
- e. Does an evaluation strategy, which shows how the implementation of learning technology has benefited students’ learning, exist?
- f. Does the institution have any effective means to disseminate the courseware locally and across the institution?
- g. Does a clear idea exist of how the courseware (or in a wider context, the curriculum) is to be managed (day-to-day, semester, work year) within the department or institution?

2. Learner Profile Analysis

- Identify the target audience
Secondary 2 students (academic)
- Identify the target audiences’ learning needs
Most students only aspire to pass examinations (or to get good grades). However, the course should be structured in a way to attempt to cultivate the development of attitudes. Science experiences in the secondary level can cultivate an interest in and love for the subject which could continue to grow long after students have left school.

- Identify and categorize target audiences into their different learning styles
As the students are streamed from secondary 1, the profile of the students is fixed. Therefore, the students should be asked to take the learning preference analysis test. Such a test enables the teacher to categorize the students into different learning styles and adopt their teaching plan according to their dominant learning style.
- Perform course-to-learner profile matching
If the course curriculum has more than one learning content presentation (in terms of learning preference), the system will perform the course-to-learner profile matching.

3. Pedagogy Analysis

- 4-Ts Theory
- Type of learning events determination

4-Ts Theory

a. Target

Goal: Improve student proficiency by a certain percentage (in terms of examination scores)

Other typical goals might include:

- Enhance student understanding (more as a supplement to traditional learning)
- Provide interactive materials (to make the lesson more interesting and to cultivate interest)
- Introduce the students to e-learning or computer aided learning
- Provide an example of computer-based learning
- As a contingency plan (e.g. SARS period, Foot and Mouth Disease)

Motivation: educational-centric

Entrance Pre-requisites:

- Primary 4 Science: Energy (partial)

- Primary 5 EM1/2/3: Systems
- Secondary 1 science

Non-coverage (exclude) List:

- Physical connection of electric circuit
- Calculation of electric circuit parameters (i.e. voltage, current, impedance, power)
- Working principle underlying electric circuit parameters (i.e. electrons flow)

b. Training

Identification of Prior Knowledge

Identification: Primary 5 Science: Systems

Concepts taught:

- o Recognize that an electric circuit consisting of an energy source and other circuit components forms an electrical system
- o Show an understanding that a current can only flow through a closed circuit
- o Recognize that (1) dry cells / battery provides energy in a closed circuit, (2) current transports energy from the dry cells / battery to the bulb and (3) a switch can be used to break or close a circuit
- o Construct simple circuits from circuit diagrams
- o Infer that components of an electrical system affect one another
- o Identify electrical conductors and electrical insulators
- o Infer that good conductors of electricity are generally good conductors of heat
- o Show an awareness of the need for proper use and handling of electricity
- o Show an awareness of the need to conserve electrical energy

Activation of Prior Knowledge

- Introduce diagrams depicting the components of electric circuits i.e. dry cells / battery, wires, bulbs, switches

- Present real life examples of electricity usage (i.e. for transport (MRT), for cooking, etc.)
- Let the students construct simple electric circuits to develop concepts and understanding through manipulation of the various components
- Use interactive circuit whereby students can manipulate the variables in the circuit to control number of dry cells / battery / bulbs or arrangement of dry cells / battery to infer how current affects the brightness of bulbs
- Use quiz at the end of the activation phase to gauge how much students remember

Identify the nature of the Training Tasks

Depending on the time constraint and the expertise of the teacher, two or more of the following training tasks can be chosen as the theme for the design.

- Collaborative / Individualistic
- Analytical / Passive
- Divergence / Convergence
- Discussion / Questioning
- Project / Assignment
- Demonstration / Practice
- Lecture / Case Study / Role Playing

Identify the Training Materials

Typically, the main subject matter and the supplementary subject matter presentation style, cognitive level and the type and amount of interactive multimedia are considered. The full contents need not be formulated here. Only the key concepts and their teaching methods are devised.

Example of key concepts

- Electric circuits – Use diagrams of circuit board to teach

Identify the Training Support

- Subject Matter Sequencing
 - Chronological

- Cognitive level
- Knowledge Gap (only if profile of learner is available)
- Taxonomic (based on structure)
- Problem / Case study-centric

To teach the science syllabus, the subject matter sequencing usually employed is the cognitive level sequencing. This means that the subject matter will increase in complexity as the course progresses. Besides adding in more complex theories, practice items and examples, higher level of learning skills (following bloom's taxonomy) such as (simple) synthesis and evaluation competency should also be catered.

- Scaffolding Means

Refer to scaffolding framework (Chapter 3, Table 3.6)

- Degree of Abstraction

The degree of abstraction chosen is the normal level of abstraction. This level is used mainly in the academic sector for imparting the basic theory. It lays the foundation for all the degree of abstraction by covering all the main learning concepts that is required for each course.

Other levels of abstraction include the overview and the expert degree of abstraction.

- Degree of Complexity

This will follow MOE level of complexity for secondary school students.

- Degree of student independence

This will follow the proposed scaffolding framework and start off with the most amount of student support. The amount of support will fade as the student gains competency.

- Pace of learning

This will follow MOE recommended coverage of course syllabus. However, as face-to-face interaction is minimized, an appropriate amount of interaction

time (in terms of group project and assignment) should be factored in to provide an all round development of the student.

c. Transfer

Pre-assessment

A pre-course assessment will be used to gauge the students' level of understanding before the commencement of the course. Besides using the result from the assessment to assess if learning has taken place (by comparing the pre-assessment results with the post-assessment results), this pre-course assessment also serves the purpose of determining where the student stands in terms of the course content. For example, if a high percentage of students fail to grasp a certain pre-requisite concept, the teacher can include the pre-requisite concept as part of the training materials.

Typical questions:

- Provide a diagram of the electric circuit components (cells, battery, power source) and ask the students to identify them
- List down the essential components of an electrical system
- Identify open and close circuit
- Infer how the components of an electrical system affect one another in terms of observation (i.e. (1) increase (2) decrease (3) no change (4) wrong connection)
- Identify diagrammatic and symbolic form of electrical conductors and electrical insulators

Practice

The practice items cover most areas that are present in the pre- and post-assessment. However, in addition to those assessment items, practice items also include projects, discussions and interactive circuitry exercise.

Examples:

- List some electrical appliances in your home and the energy changes that take place in each of them
- What did people in Singapore use before homes had electricity

- Make a list of things in your home that use electricity from the mains and from electric cells
- Analyze the circuit to see which bulb does not light up
- Rearrange the circuit to get the bulbs to light up

Post-assessment

Typical questions:

- How the arranging of resistors in series or parallel affect the resistance of a circuit
- Test the heating, magnetic and chemical effects of an electric current and their applications
- Test the meaning of current and voltage, their units and how to measure them
- How to draw and interpret the basics of electric circuits

Remediation

Depending on the assessment results of the student's performance and the analysis of the misconceptions or doubts that each particular student has, particular sections of course content could be revisited. While the theoretical explanation/presentation format might remain (or change to cater to another dominant learning style), the practice items will however, not be repeated. The selection of the practice items can either be randomized or customized by the teacher. Usually, if the grade of the student falls below the accepted rate by a significant margin, the teacher should arrange a face-to-face meeting to find the problem and to customize the e-learning solution.

Enhancement – introduction to next level of understanding

This effectively acts as a prelude to the next course. Depending on the nature of the learning materials for the next course, different enhancement items are introduced. The enhancement items typically are open-ended in the sense that it invites the student to reflect on 'existing' knowledge.

Examples:

- Two friends were talking about electric cells and has a hypothesis
A: “The size D cell is larger than the size AA cell, so it will make the bulb brighter”
B: “The size D cell will make the bulb shine as bright as the size AA size but will last longer”
Question: Who is correct? Design an experiment to find out.
- A factory that collects scrap metal wants an effective method to separate iron and steel scrap from the other metallic scrap on a conveyor belt. Design a machine to do this.

d. Transformation

Assess the accountability of the teaching

This compares existing knowledge (from pre-assessment) with the new knowledge (from post-assessment) and gauges if there is any progress.

Feedback – both learners and teachers

Feedback can take the form of a survey (conducted at the end of the lesson) or by email (as and when the student feels it is necessary).

Revision

Revision items are usually formed by an overview of key theories coupled with recap questions.

Example:

Theory

An **electromagnet** is formed when a coil of wire is wound around a piece of iron. When a current flows through the coil, it behaves like a bar magnet, but when the current stops flowing through the coil, the electromagnet loses its magnetism.

Question:

What are the ways that an electromagnet can be made stronger?

Type of learning events determination

This is typically factual learning.

4. Course Profile Analysis

- Identify the overall (high level) educational Goal
To teach the basics of electrical circuits
- Identify the conceptual scope
This defines the boundaries of the application domain. For a start, the course targeted learning objectives and learning concepts to be imparted, can be used as a gauge to determine the amount and level of details to be included in a particular course. The targeted audience's expertise and academic knowledge will also play a part in determining the scope.

The conceptual scope will follow MOE's guideline on the amount and level of detail.

- Identify the application domain
Academic
- Identify the type of performance augmentation
To comprehend the learning concepts and to appreciate the operation of real world electrical appliances

Design Phase:

1. Performance Design

- Identify all the measurable learning and performance objectives
- Determine the assessment guidelines

Measurable Learning Objectives

- How to draw and interpret circuit diagrams involving a variety of electrical components
- The meaning of current and voltage, their units and how to measure them
- The meaning of resistance, its unit and how its variation affects the current in the circuit

- The arranging of resistors in series/parallel effects the resistance in the circuit
- The heating, magnetic and chemical effects of an electric current and their appliances

Performance Objectives

- Appreciate the working principles behind electrical appliances

Assessment Objectives

Following MOE's assessment objectives:

The assessment objectives for this subject are classified into three main categories: knowledge with understanding; handling, applying and communicating information (science process skills and thinking skills); and exploration and investigation.

(I) Knowledge with understanding

Students should be able to demonstrate knowledge and understanding of

- scientific phenomena, facts, concepts and principles;
- scientific vocabulary, terminology and conventions;
- scientific instruments and apparatus including techniques and aspects of safety;
- scientific and technological applications.

(II) Handling, applying and communicating information (science process skills and thinking skills)

Students should be able to use visual, aural and written (including symbolic, diagrammatic, graphical and numerical) information to:

- locate, select, organize and present relevant information from a variety of sources;
- transpose information from one form to another;
- process numerical and qualitative data;
- identify patterns, report trends and draw inferences;
- present reasoned explanations for phenomena, patterns and relationships;
- make predictions and hypotheses and deduce relationships;
- apply knowledge, including principles, to novel situations and problems;
- communicate ideas and observations effectively.

(III) Exploration and investigation

Students should be able to:

- follow and carry out instructions accurately and safely;
- identify the problem and then plan and carry out an investigation to solve the problem;
- use and organize techniques, apparatus and materials effectively and safely;
- observe, measure and record accurately;
- interpret, evaluate and establish validity of the observations and experimental data;
- evaluate methods and suggest possible improvements.

Based on the design assessment objectives and following Bloom's Taxonomy, the following is devised.

Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
30%	30%	20%	10%	5%	5%

2. Scaffolding Design

Scaffolding is designed to promote teaching strategies across the institution that respects the concept that *all* students have different ways of learning.

- Identify the learning context that the concept resides in
Academic Context
- Scaffolding Goal Design

Formulate:

- Scaffolding Aim
- Scaffolding Problem

Scaffolding Aim

The teacher must first identify the curriculum goals and select the context whereby the learning concept will reside in. He must be knowledgeable of the content and be sensitive to the learners (e.g. aware of the learners' background knowledge and misconceptions) to determine if they are making progress. Using previous curriculum records of students and based on the teacher's past

experience in teaching the same course, the following areas are good candidates for scaffolding:

- Use scaffolding when the class consists of students with a significant variance in their cognitive levels. Scaffolding can be used for all types of students. For students who are far below the recommended standard, scaffolding is used to provide support for under-developed or deficient skills. For above average students, it is important to appeal to their higher-level reasoning or insights, while allowing scaffolding or compensation for the problem areas involving perceptual processing. Uneven abilities will occur in all classes, and is of concern when one's primary teaching techniques is scaffolding (although it should not be the primary style). Variation is more exaggerated and pronounced for students with extreme cognitive strengths and weaknesses, who are on the other end of the spectrum. The theory of Multiple Intelligences explains this fact of different abilities [Henry, 2002].
- Use scaffolding to engage students in research work and learning as scaffolding facilitates organization of and focus for students' research [McKenzie, 1999]. The structure and clearly defined expectations are the most important components of scaffolding in this context. The teachers provide clarity and support but the students construct the final result through their research.
- Use scaffolding to teach scientific inquiry and experimentation (external representations – graphs, tables, etc.). It is found that the use of external representations, representational scaffolds, can serve as an effective strategy for teaching these scientific skills. E.g. *In one study the instructional goal was to teach fourth graders valid experimentation skills. During the first part of the study, a teacher-specified table of variables was the only scaffold provided. Students had to select the appropriate variable related to their experiment. The results of this part of the study led to the conclusion that the "... use of the pre-developed table representation may have helped students abstract the overall structure of*

the experiment and thus aided their understanding of the design..." [Toth, Results and Discussion section, para. 1]. *The teacher-designed table helped focus the learners' thinking on only those items that were important for the task. Additionally through the use of the table it became obvious to the students if they had omitted an important variable from their experiment. This helped the students learn what things must be considered when designing an experiment* [Toth, n.d)).

- Use scaffolding instruction for guiding students to be independent and to promote the cultivation of self-regulated competence of skills [McKenzie, 1999].
- Use scaffolding to develop thinking skills and to facilitate advanced thinking. For example, break a complex task into meaningful chunks and ask students to undertake specific sub-tasks. By scaffolding and stepping through the tasks, the student can gain the kind of thinking process that more expert students would typically use. This also follows the constructivism perspective that when students need to understand a more complex or sophisticated topic, it doesn't help to serve them simplified truths, boiled down examples, or step-by-step formulas. What they need are many examples with lots of information and opinions on the topic through which they will sift until they have constructed an understanding that not only connects to their individual prior knowledge, but also builds new schema that will be refined when students encounter the topic again in the future. **Until e-learning, this kind of activity was very difficult for the average teacher to create because collecting such a breadth of resources was next to impossible.**
- Use scaffolding for problem-based learning [Greening, no date].
- Use scaffolding when the students are classified into a vast varieties of learning styles and preferences
- According to Education World, Scaffolding can work in a number of ways depending on the type of instruction being provided. For subject areas in which knowledge is "fixed," teachers can provide all the needed support

beforehand. For subject areas in which knowledge is developmental, more discussion and ongoing guidance may be necessary [Gay, 2002].

- Scaffolding can apply to any number of situations within the classroom and in every day life. Scaffolding is appropriate whenever there is a discrepancy between what a child knows or is capable of, and the developmental step-up to a more expert-like level of knowledge. This method is appropriate well into adulthood, as new skills are learned which are more efficiently obtained through the modeling of others rather than trial and error, or self-guided learning.

Typical example of scaffolding aim:

- To hand over the knowledge and control to the students
- If the teaching aim is to equip pupils with a lexicon, and guide their responses towards the explanative and analytical, then the teaching plans should incorporate scaffolding as part of the teaching plan
- To enhance the learning experience of the students. Typically, students who are able to master a task while others struggle, may find themselves frustrated. This frustration comes as a result of the student having to wait for the teacher to redirect attention to the progress of the lesson. Hence, a possible solution is to have the students to assist other students in the scaffolding process.

Scaffolding Problem

After the initial identification of the scaffolding aim, the teachers will identify possible problems that could occur when using scaffolding.

Typical implementation problems are:

- Network-intensive/delivery problems
- Time constraints
- Multi-media presentation problems
- Technology-constraints

Typical learning problems are:

- Which part of the learning content can be scaffolded
- What percentage of the learning content is to be scaffolded
- Are students skeptical about being categorized into different scaffolding levels
- How to efficiently determine the scaffolding levels of the students

■ Scaffolding Tasks Design

Formulate Tasks to address:

- Activation of prior-knowledge
- Encoding specificity – resemblance of problem to learning concepts
- Elaboration of knowledge application

Activation of prior-knowledge

Similar to the task conducted at 4T: Training except that this task only looks at one learning concept and traces deeper into the student's cognitive structure.

Encoding specificity

Instead of purely identifying the prior knowledge, this task links up the prior knowledge to the current scaffolding concept. The linkage will be formed through the usage of a concept map that includes relations such as essential pre-requisites, good to know pre-requisites, augment post-requisites, utilize pre-requisites and associated courses.

Elaboration of knowledge application

The aim of the elaboration of the knowledge application is to produce the knowledge constituent model and task hierarchy. It describes how the application semantics can be decomposed and how each knowledge constituent is elaborated according to pedagogical logic. Essentially, the knowledge constituent model consists of the knowledge breakdown and the elaboration of each constituent. The knowledge breakdown seeks to modularize knowledge into a hierarchical structure and allocate them into different course offerings. Based on the

description of each knowledge constituent, the adopted learning/teaching logic must be stated. The modularized course offering acts as a first cut to the RLO/RIO identification which will be conducted in section 3 of the design phase.

- **Learner Considerations**

Formulate Tasks to address:

- Identify the learner's goal through the process of needs analysis
- Knowledge Gap Analysis
 - Assess learner's current knowledge point (Assess what the learner is currently able to do without help)
 - Assess learner's desired knowledge point (Find the desired performance)
 - Determine the level of learner support to design
(While all 3 levels should be designed, the most appropriate level should be determined and designed first and used as a guide to spearhead the design)
 - Knowledge Gap Mapping
(Analyze what has to be achieved to move the learner from the current knowledge point to his desired goal)

Needs Analysis

Before the scaffolding aim can be formulated, the teacher must first conduct the students' needs analysis. Knowledge of student needs and the analysis of why they fail to achieve their needs are essential pre-requisite activities prior to the formulation of the scaffolding aim. It is only with the clear scaffolding aim in mind that the teacher can effectively plan the teaching routes and methods that are truly learner-centric.

Knowledge Gap Analysis

Based on the results from the needs analysis, the student's current knowledge point can be acquired. Next, the targeted knowledge point can also be easily

known as this is the main scaffolding concept. With these two points known, the level of learner support (based on the scaffolding table) can then be assessed.

- Specify what is necessary to narrow or eliminate the knowledge gap
This effectively investigates the number of learning concepts to be mastered.
- Specify what is the appropriate instructional strategy to scaffolding the design
Select between cognitive / situated restructuring
- Determine how to measure success or failure
This sets the assessment criteria.

3. Course Structure Design (Extrinsic Properties)

- RLO Identification (Identify the individual modules that make up a course)
Formulate:
 - Educational Goal
 - Cognitive Level
 - Type of content presentation
 - Associated Keywords
- RIO Identification (Identify the individual topics that make up a module)
Formulate:
 - Educational Objective
 - RIO type (Concept/Fact/Procedure/Process/Principle)
 - Associated Keywords
- Course Tree Hierarchy
- Course concept map
- RLO/RIO search for reused/repurposed
- Verification of course structure
- Verification of RLO/RIO status (develop/reuse/repurpose)

Section 3 of the design phase looks at categorizing the knowledge offerings into different RLO/RIO modules and housing them in a course concept map. It will act upon the results from the elaboration of the knowledge application.

References:

[Gay, 2002] Gay, G. (2002). Lesson Design - Constructing Knowledge and Scaffolding Instruction. Retrieved November 02, 2005, from <http://snow.utoronto.ca/Learn2/greg/4294/scffldng.htm>

[Greening, n. d.] Greening, T (no date). Scaffolding for success in PBL. Med Educ Online [serial online] 1998; 3, 4. Retrieved November 1, 2005, from <http://www.utmb.edu/meo/>

[Henry, 2002] Henry, L. (2002). Educational Concept of Scaffolding. Adolescent Learning & Development.

[McKenzie, 1999] McKenzie, J. (1999). Scaffolding for Success. [Electronic version] Beyond Technology, Questioning, Research and the Information Literate School Community. Retrieved October 1, 2005, from <http://fno.org/dec99/scaffold.html>

[Perelman, 1998] Perelman et al. (1998). The Mayfield Handbook of Technical & Scientific Writing. Mayfield Publishing Company, 1998.

[Toth, n. d.] Toth, E. E. (no date). Representational scaffolding during scientific inquiry: interpretive and expressive use of inscriptions in classroom learning. Retrieved October 1, 2005, from <http://www.cis.upenn.edu/~ircs/cogsci2000/PRCDNGS/SPRCDNGS/posters/toth.pdf>

APPENDIX B

SCAFFOLDING FRAMEWORK – ILLUSTRATIONS

B1 Scaffolding

The scaffolding methodology is divided into three levels and requires the learning content to be divided into at least three sections for the initial classification. The number of sections the learning content is to be divided into depends on:

1. The complexity of the learning content in general

The more complex the learning content, the more sections the learning content should be separated to simplify the learning curve per course offering.

2. The estimated learning time required

The longer the estimated learning time for the whole courseware, the more time there is for the teacher to include a variety of pedagogical/scaffolding methods. Hence, a long course duration will allow the learning content to be separated into more sections.

3. The ease of separating the whole courseware into sections

The ease of separating the courseware will depend on the number of learning concepts to be learned. If there are different learning concepts, the learning content should be separated into more sections.

4. The dependency between different learning concepts

Learning concepts that are inter-related should be offered in a single module offering. Hence, as the dependency between the learning concepts increases, the fewer the number of sections the learning content should be separated into.

5. The dominant/preferred teaching method/strategy

If there are different teaching methods for different learning concepts, each section of the learning content should be taught by one teaching method. Hence, the greater the diversity in the teaching method, the more sections the learning content should be separated into.

Each section of learning content will typically start off with the impartation of learning contents (typically consists of teaching approximately 70~100% of the underlying principles behind the learning content), followed by a guided (step by step) practice/demonstration/illustration/application of the learning content. Next, the student

will be asked to do some hands-on practice (unguided) before they are assessed for possible promotion into the next scaffolding level / content. As the scaffolding level increases, the amount of scaffolds or student support will decrease.

For students who pass the first level of assessment and meet the pre-defined criteria, they will be promoted to the second level of scaffolding where the student support is minimized. At level 2, the same sequence of learning (impartation → guided practice → unguided practice → assessment) will be conducted. However, at this level, the amount of direct teacher-teaching is minimized in the sense that only about 40~70% of the learning content will be covered in detail. The student will also be asked to investigate/infer the rest of the learning contents through guided hints. Pair and group discussions are usually conducted (to mitigate the loss of scaffolds) in this level. Typically, hints will take the form of reference books, open-ended questions and links to external learning materials. Fading takes place and the amount of guided practice is minimized. However, the students are required to take a greater amount of unguided practice before they can be assessed for promotion into the next level.

Similarly, students who passed the second level of assessment will be promoted to the third and last level of the scaffolding hierarchy. Here, student support is kept at a minimal (if not none). While some key concepts are still covered, most of the underlying concepts have to be inferred. The unguided practice items will also take up the bulk of the learning time as it is usually scenario-based and covers the use of a combination of learning concepts. The full learning initiative lies at the hands of the students. Teacher's help is usually asynchronous and negligible. However, the student is free to seek any form of external help (from fellow peers or friends).

However, in the event that the student cannot cope with the reduced scaffolds, additional help in the form of both peers and teacher will be accorded. The level of scaffolding that the student is currently in will not decrease in hierarchy level even if the student fails in the assessment level. While the students may encounter greater difficulties with the reduced scaffolds and also experience dissatisfaction/disillusioned with their learning abilities, we believe that it is only through such exposure that the student will pick up the

necessary skills that we believe are so important in their character shaping and learning. However, should the student constantly fail to display the expected amount of expertise (i.e. fails the assessment for more than three times), the teacher will be notified. He will then have to assess the learner and decide if the current scaffolding level is suitable for the learner. This is important. Many learners drop out of e-learning courses due to the lack of motivation that is garnered from a constant experience of failure and the lack of personalized support. Hence, it is crucial that such appalling learning experiences be avoided.

Illustration

The scaffolding assessment is illustrated using three sections of learning content. All learners will start from scaffolding level 1 and proceed according to the structure of the course. Assessment will be conducted at the end of each learning content section. Assessment will be based on the understanding and application of learning concepts covered. If the student fails the assessment, they will be required to revisit the section again. Passing the assessment will allow the student to take on the next section of learning content. The level of scaffolding for the next section will depend on the assessment grade; i.e. students who displayed greater competency (grade > 75%) in a particular section will be promoted to the next level of scaffolding while students will only be demoted to a lower level of scaffolding if he (1) failed the assessment for more than three times and (2) the teacher has suggested demotion. The scaffolding level will however remain the same for students who passed the assessment but did not attained > 75% of concept understanding.

The flow chart depicts how the scaffolds will change for different levels of scaffoldings. However, for illustration purposes, the flow chart only depicts the scaffold changes for students who passed the assessment.

Appendix B: Illustration of Scaffolding Framework

Appendix B: Illustration of Scaffolding Framework

Appendix B: Illustration of Scaffolding Framework

Appendix B: Illustration of Scaffolding Framework

Appendix B: Illustration of Scaffolding Framework

APPENDIX C

VECTOR MODEL – BASIC CONCEPTS AND CLUSTERING ANALOGY

C1 Basic Concept of Vector Model

The classic model of Information Retrieval (IR) considers document representation by indexing. Each document can be represented by a set of representative keywords called index terms. Index terms are contextually important words that can be used to summarize, describe or index a document. These terms (either individually or as a collection) have semantics that are able to fully depict the objective of the document. They can be extracted automatically either from the documents content or from manually assigned descriptor terms (i.e. metadata). In general, index terms are mainly nouns or stemmed words. Index terms have varying relevance when used in a certain application context. Such semantics are captured through the assignment of numerical weights to each index term in a document.

The following assumptions are commonly assumed in the context of IR:

1. All documents can be represented as a set of index terms.
2. The index terms are keywords (mostly nouns) that are extracted from either the metadata or extracted from the document to summarize its contents.
3. The importance of individual index terms to a document is different; not all extracted index terms carry the same importance in summarizing the contents of the document. The importance of an index term can be represented as a weight.

A naive though important assumption is made that the terms in a document are mutually independent. The term independence means that the appearance of one term in a document is unrelated to the appearance of another term. This assumption is made to simplify the computation of the index term weights and also to allow faster ranking computation. However, this is often inaccurate. For example, given a network security course under the underlying assumption, the term of firewall and gatekeeper are mutually independent; that is $P(\frac{firewall}{gatekeeper}) = P(firewall)$. However, logically, if the term of gatekeeper appears, the probability of the term firewall appearing will be higher than if the term does not appear. Nonetheless, it has been demonstrated that performance is still quite good under this assumption [Baeza-Yates, 1999].

Mathematically, documents and their query are represented as vectors in a t -dimensional space. Let $K: \{k_i\}$ be a set of index terms, $D: \{d_j\}$ be a set of documents, $Q: \{q_k\}$ be a set of queries, and $R: D \times Q \rightarrow R(Q)$, where $R(Q) \subseteq \mathbb{R}$ is a function that assigns each document and each query with a real number that represents the ranking or relevance of the document with respect to the query.

[Definition C1] Let t represent the number of terms in the entire system and k_i represent an index term where $K = \{k_1, k_2, k_3, \dots, k_t\}$ is the full set of index terms.

[Definition C2] Let w_{ij} be a real number that represents the ranking or relevance of the document with respect to a particular query criterion. $w_{ij} \geq 0$ indicates that the weight is associated with a particular index term k_i in a document d_j ; i.e. the pair $(k_i, d_j).w_{ij} = 0$ when k_i does not appear in d_j .

[Definition C3] Let D represent a set of documents and d_j represent a single document where $D = \{d_1, d_2, d_3, \dots, d_t\}$ is the full set of documents. A document d_j is associated with the index term vector and is a set of index terms with weights; a weight expresses the relative importance or relevance of the term with respect to the document.

$$\vec{d_j} = (w_{ij}, w_{2j}, \dots, w_{tj})$$

[Definition C4] Let q represent a single learner's query. The query q is a set of query terms with weights; a weight expresses the relative importance or relevance of the term with respect to the learning needs of the learner.

$$\vec{q} = (w_{iq}, w_{2q}, \dots, w_{tq})$$

The Vector Model is adopted to compute the degree of similarity between the learner's query and the course document. For each document d_j and query q , a similarity measure $\text{sim}(d_j, q)$ measures the degree of similarity between the document and the query. The similarity measure will increase as similarity grows (i.e. 0 indicates total dissimilarity and 1 indicates perfect similarity or match). Using the similarity measure, the entire collection

of course documents in the specified domain will be ranked in response to the learner's query.

The degree of similarity will be used as a first cut to rank the retrieved documents. Documents, whose similarity to the query exceeds a certain threshold value, v , will be retrieved for further processing. As queries are analogous to documents, the same similarity measure can be used to measure (1) document-query similarity (used in searching), (2) document-document similarity (used in document clustering), (3) term-term similarity (used in term clustering, statistical thesauri) and (4) query-query similarity (for agent reasoning and inference). As the similarity measure employed will only be used to rank the search results, only the document-query similarity measure will be discussed.

The similarity measure has the following characteristics:

- Non-negative: $\text{sim}(d_j, q) \geq 0$
- Symmetric: $\text{sim}(d_j, d_j) = \text{sim}(d_j, d_j)$
- $\text{sim}(d_j, q) = 0$ indicates perfect dissimilarity
- A higher value of $\text{sim}(d_j, q)$ indicates higher similarity
- The value of $\text{sim}(d_j, q)$ is often normalized to values in the range of $\{0 - 1\}$

The document d_j and user query q can be represented as a t -dimensional vector as shown in figure 4.1. The Vector Model evaluates the degree of similarity of d_j with respect to q in terms of the correlation between the vector $\vec{d_j}$ and $\vec{q}$. This correlation can be quantified by the cosine of the angle between these two vectors.

Given: Document: $\vec{d_j} = (w_{1j}, w_{2j}, \dots, w_{tj})$

Query: $\vec{q} = (w_{1q}, w_{2q}, \dots, w_{tq})$.

The cosine similarity measure:

$$\text{sim}(d_i, q) = \frac{\vec{d_i} \cdot \vec{q}}{|\vec{d_i}| \times |\vec{q}|} = \frac{\sum_{i=1}^t w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^t w_{i,j}^2} \times \sqrt{\sum_{i=1}^t w_{i,q}^2}} \quad (\text{eq. C1})$$

where $\left| \vec{d}_i \right|$ and $\left| \vec{q} \right|$ is the normalization value of the document and query vector.

Figure C1 The cosine of θ between d_j and q

Figure C1 shows the graphical view of $\text{sim}(d_j, q)$ which measures the cosine of the angle between the query and the document vector. As shown by the figure, as the cosine value approaches 1 (or the two vectors become coincident), the document and the query represents the same concept. On the other hand, as the cosine value approaches 0 (or the vectors become orthogonal), the document and the query represents entirely unrelated concepts. Thus, given that $w_{ij} \geq 0$ and $w_{iq} \geq 0$, $\text{sim}(d_j, q)$ will varies from 0 to 1.

C2 Basic Principle of the Clustering Analogy

Mathematically, the basic principles of the clustering analogy can be summarized as follows:

[Definition C5] Let documents be seen as a collection of C objects and query be a vague description of requirements A where A is a subset of C . Thus, for IR, we aim to partition C into sets of A and $\neg A$ where A contains object features that describe the query and $\neg A$ contains object features that best differentiate A from $\neg A$ (Figure C2).

Figure C2 Complete Query Match

Figure C3 Partial Query Match

In the Vector Model, the tf-idf (term frequency - inverse document frequency factor) scheme is employed. The term frequency factor is used to quantify the intra-cluster similarity; if a term t_i appears often in a document d_i , then a query q containing t_i should retrieve that document d_i . The term frequency factor, tf , is given by $f_{i,j}$:

$$f_{i,j} = \frac{freq_{i,j}}{\max_j freq_{i,j}} \quad (\text{eq. C2})$$

Where $f_{i,j}$ is the normalized frequency of terms k_i in document d_i and $\max_j$ is the maximum frequency that can be computed over all the terms which exist in document d_i .

The term inverse document frequency factor, idf , is given by idf_i :

$$idf_i = \log\left(1 + \frac{N}{n_i}\right) \text{ or } idf_i = \log\left(\frac{N - n_i}{n_i}\right) \quad (\text{eq. C3})$$

Where N is number of documents in a collection and n_t is the number of documents containing the term.

The best known term-weighting schemes use weights which are given by

$$w_{i,j} = f_{i,j} \times idf_i \quad (\text{eq. C4})$$

APPENDIX D

EXPERIMENTAL DATA

No.	Course	Module
01	Software Engineering	An Introduction to Software Architecture
02		Overview of Software Engineering I
03		Overview of Software Engineering II
04		Overview of Software Engineering III
05		Object-Oriented Concepts and Notations I
06		Object-Oriented Concepts and Notations II
07		Object-Oriented Software Development Process
08		Some Generic Object-Oriented Design
09		Object-Oriented Analysis and Design Using UML - Case Studies
10		Measuring Object-Oriented Design
11		Emerging Topics in Software Engineering
12	System Engineering (and Graph Theory)	Design of System Architectures
13		Systems Engineering
14		System Concepts
15		Sociotechnical Systems
16		Using the WinWin Spiral Model: A Case Study
17		Introduction to Graph Theory
18		Graphs- Basics
19		Graph Applications I
20		Graph Applications II
21	Networking and Databases	Network Design: ISO/OSI reference model
22		Network Design: Local Area Networks
23		Network Design: Internet TCP/IP and Internetworking
24		Network Design: Industrial Networking - Industrial LAN
25		Network Design: Internet Application and Issues
26		Network Design: Web Application Development Tools
27		Network Design: Industrial Networking - Serial Networking
28		Database and Middleware: Database and Middleware Introduction
29		Database and Middleware: Entity-Relationship models
30	Knowledge Engineering	Knowledge Engineering Overview I
31		Knowledge Engineering Overview II
32		Agent Introduction I: Why, When, and Where to Use Software Agents
33		Agent Introduction II: Agent-Based Systems for Intelligent Manufacturing: A State-of-the-Art Survey
34		Concurrent Knowledge Processing Techniques I
35		Concurrent Knowledge Processing Techniques II
36		Autonomous Agents/Actors I
37		Autonomous Agents/Actors II
38		Knowledge Processing Techniques
39		Machine Learning
40		Tools for the Development of Knowledge Processing Systems
41		Evolutionary Computation

42		Data Mining and Knowledge Discovery
43		Autonomous agent - AI Planning
44		Sub-Symbolic Learning

$$\text{Formula: } \text{sim}(d, q) = \frac{\vec{d}_i \cdot \vec{q}}{\|\vec{d}_i\| \times \|\vec{q}\|} = \frac{\sum_{i=1}^t \text{Vector}_{\text{KeyCon}} \times \text{Vector}_{\text{freq}}}{\sqrt{\sum_{i=1}^t \text{Vector}_{\text{KeyCon}}^2} \times \sqrt{\sum_{i=1}^t \text{Vector}_{\text{freq}}^2}}$$

Software Engineering

An Introduction to Software Architecture				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	architecture	165	system	198
02	design	65	architecture	165
03	architectural styles	24	data	153
04	case studies	20	software	151
05			component	75
06			control	73
07			model	69
08			level	68
09			design	65
10			blackboard	50

$$\text{sim}(d, q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0.025, 0.0, 0.0, 0.0, 0.0, 0.25, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.025^2 + 0.25^2}} \\ = 0.396$$

Overview of Software Engineering I				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	software engineering	39	software	81
02	software product	5	engineering	42
03	software process	2	software engineering	39
04	myths	5	system	19
05	engineering principles	3	cost	18
06			process	17
07			development	16
08			requirement	12
09			product	8
10			design	7

Appendix D: Experimental Data

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0.2, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.230 \end{aligned}$$

Overview of Software Engineering II				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	process model	5	system	45
02	prototype	16	software	31
03	prototyping language	1	engineering	24
04			software engineering	23
05			development	19
06			process	17
07			prototype	16
08			requirement	11
09			evolutionary	7
10			process model	5

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0.33, 0, 0, 0.33)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.238 \end{aligned}$$

Overview of Software Engineering III				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	paradigm	7	software	23
02	requirement engineering	1	requirement	21
03	Modeling	11	engineering	15
04	specification	6	modeling	11
05			functional	9
06			system	8
07			paradigm	7
08			process	7
09			design	7
10			specification	6

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0.25, 0, 0, 0.25, 0, 0, 0.25)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.432 \end{aligned}$$

Object-Oriented Concepts and Notations I				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	association	61	object	99
02	inheritance	31	class	89
03	generalization	30	association	61
04	aggregation	22	oriented	57

Appendix D: Experimental Data

05	constraint	25	notation	55
06			instance	41
07			inheritance	31
08			generalization	30
09			constraint	25
10			function	23

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0.2, 0, 0, 0, 0.2, 0.2, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.490 \end{aligned}$$

Object-Oriented Concepts and Notations II				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	dynamic modeling	2	case	51
02	use case	50	use case	50
03	interaction diagram	8	state	39
04	State machine	9	notation	27
05	activity diagram	9	object	27
06			concept	24
07			system	20
08			flow	11
09			transition	11
10			behavior	10

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.2, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.260 \end{aligned}$$

Object-Oriented Software Development Process				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	unified process	5	model	51
02	use case	39	use case	39
03	architecture-centric	2	architecture	35
04	UML	1	process	34
05			development	25
06			design	18
07			implementation	16
08			analysis	14
09			system	13
10			unified	7

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.25, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.325 \end{aligned}$$

Some Generic Object-Oriented Design				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	persistence	12	design	49
02	dbms	18	object	46
03	relational dbms	2	generic	43
04	normalization	9	data	30
05	normal form	5	database	27
06			class	22
07			table	21
08			dbms	18
09			persistence	12
10			normalization	9

$$sim(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 2, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} = 0.144$$

Object-Oriented Analysis and Design Using UML - Case Studies				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	OOAD	43	case	98
02	case study	88	case study	88
03	UML	46	analysis	58
04	requirement analysis	12	design	53
05			UML	46
06			OOAD	43
07			requirement	21
08			requirement analysis	12
09			class	12
10			specification	9

$$sim(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.25, 0, 0, 0.25, 0.25, 0, 0.25, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} = 0.401$$

Measuring Object-Oriented Design				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	metric	21	class	57
02	evaluation criteria	1	criterion	42
03	Weyuker Measurement Principle	1	design	39
04			method	38
05			measure	30
06			object	28
07			metric	21
08			complexity	14

Appendix D: Experimental Data

09			evaluation	12
10			function	6

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0.33, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.191 \end{aligned}$$

Emerging Topics in Software Engineering				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	emerging topic	72	software	106
02	architectural style	15	emerging topic	72
03	architecture description language	4	system	65
04	component based	5	architecture	60
05			engineering	53
06			object	29
07			component	28
08			style	27
09			data	16
10			transformation	16

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.25, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.325 \end{aligned}$$

System Engineering (and Graph Theory)

Design of System Architectures				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	system architecture	85	system	228
02	design	191	design	191
03	heuristic	23	architecture	182
04	pattern	75	product	173
05	product family	30	part	95
06			system architecture	85
07			family	78
08			pattern	75
09			manufacturing	65
10			flow	45

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.25, 0, 0, 0.25, 0, 0, 0.25, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.650 \end{aligned}$$

Systems Engineering				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency

Appendix D: Experimental Data

01	system engineering	148	system	329
02	system lifecycle	5	design	186
03	model	65	engineering	168
04	system validation	11	requirement	156
05	system verification	0	system engineering	148
06	Design Analysis	5	testing	132
07			product	119
08			lifecycle	97
09			analysis	91
10			process	91

$$sim(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0,0,0,0,0.167,0,0,0,0,0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ = 0.144$$

System Concepts				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	system concept	24	system	275
02	open system	21	function	66
03	closed system	15	part	62
04	system classification	0	environment	47
05	structure	29	property	35
06			structure	29
07			concept	28
08			system concept	24
09			layer	22
10			open system	21

$$sim(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0,0,0,0,0.2,0.2,0,0,0.2)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ = 0.260$$

Sociotechnical Systems				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	sociotechnical systems	7	work	66
02	social system	8	system	51
03	technical system	5	analysis	46
04	CIM	3	design	41
05	manufacturing	18	technical	33
06	product development	2	technology	33
07			social	32
08			process	29
09			group	29
10			manufacturing	18

Appendix D: Experimental Data

$$\text{sim}(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 0, 0.167)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ = 0.024$$

Using the WinWin Spiral Model: A Case Study				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	winwin	65	system	76
02	spiral model	30	model	67
03	cost-effectiveness	1	software	66
04	software engineering	19	winwin	65
05	UML	3	development	47
06			spiral	41
07			requirement	36
08			process	34
09			architecture	33
10			engineering	31

$$\text{sim}(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0.2, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ = 0.201$$

Introduction to Graph Theory				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	graph	156	graph	156
02	digraph	26	edge	68
03	circuit	18	vertex	55
04	cycle	28	tree	51
05	eulerian graph	1	degree	43
06	hamiltonian graph	3	connected	36
07			cycle	28
08			digraph	26
09			eulerian	23
10			hamiltonian	21

$$\text{sim}(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0.167, 0, 0, 0, 0, 0, 0.167, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ = 0.410$$

Graphs- Basics				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	component equations	10	graph	87
02	fundamental cycle	7	component	53
03	fundamental cut set	12	matrix	52
04	terminal component	8	edge	44
05	cycle law	4	circuit	42
06	vertex law	3	fundamental	37

Appendix D: Experimental Data

07	component modeling	4	system	34
08			vertex	32
09			source	24
10			terminal	23

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.000 \end{aligned}$$

Graph Applications I				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	euler trail	3	equation	33
02	graph coloring	2	flow	33
03	network flow	2	network	29
04	Max-Flow Min-Cut Concept	1	graph	22
05	physical network	2	source	19
06			fundamental	15
07			vertex	14
08			terminal	13
09			component	13
10			matrix	9

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.000 \end{aligned}$$

Graph Applications II				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	Graph Tableaus	1	branch	16
02	f-cutsets	4	chord	14
03	f-circuits	7	circuit	14
04	Graph Analysis	1	cutset	8
05			f-circuits	7
06			tableau	6
07			tree	5
08			graph	5
09			f-cutsets	4
10			matrix	4

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0.25, 0, 0, 0, 0.25, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.287 \end{aligned}$$

Networking and Databases

Network Design: ISO/OSI reference model				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	network layer	13	layer	131
02	OSI	43	network	75
03	ISO	42	OSI	43
04	reference model	8	data	43
05	network design	15	ISO	42
06			protocol	24
07			transport	18
08			link	17
09			physical	16
10			network design	15

$$sim(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 2, 0, 0, 2, 0, 0, 0, 2)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ = 0.431$$

Network Design: Local Area Networks				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	local area network	5	LAN	82
02	LAN	82	system	35
03	topology	30	token	34
04	LAN system	28	network	33
05	LAN connection	12	topology	30
06			LAN system	28
07			data	27
08			connection	26
09			ring	23
10			physical	20

$$sim(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0.2, 0, 0, 0, 0.2, 0.2, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ = 0.606$$

Network Design: Internet TCP/IP and Internetworking				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	internet	60	protocol	97
02	TCP	29	network	80
03	IP	51	application	66
04	TCP/IP	16	internet	60
05	internetworking	2	IP	51
06			layer	49
07			connection	41

Appendix D: Experimental Data

08			data	39
09			transport	35
10			TCP	29

$$\text{sim}(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0.2, 0.2, 0, 0, 0, 0, 0.2)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}}$$

$$= 0.403$$

Network Design: Industrial Networking - Industrial LAN				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	industrial networking	13	network	49
02	distributed applications	8	industrial	27
03	network design	12	interface	26
04	information backbone	10	data	25
05			manufacturing	22
06			control	21
07			distributed	14
08			design	14
09			industrial networking	13
10			network design	12

$$\text{sim}(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0.25, 0.25)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}}$$

$$= 0.107$$

Network Design: Internet Application and Issues				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	internet	21	search	25
02	web basics	1	internet	21
03	web browser	1	server	21
04	web server	8	access	10
05	web issue	1	information	10
06			computer	10
07			web server	8
08			tier	8
09			cookies	7
10			client	6

$$\text{sim}(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.2, 0, 0, 0, 0, 0.2, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}}$$

$$= 0.376$$

Network Design: Web Application Development Tools				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	internet	11	java	52

Appendix D: Experimental Data

02	web	25	element	41
03	web service	6	class	38
04	java	52	data	37
05	data presentation	1	html	30
06			language	27
07			web	25
08			model	23
09			markup	15
10			program	14

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0.2, 0, 0, 0, 0, 0.2, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.404 \end{aligned}$$

Network Design: Industrial Networking - Serial Networking				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	serial networking	1	fieldbus	67
02	industrial networking	1	control	31
03	fieldbus technology	13	layer	26
04			device	16
05			data	25
06			foundation	17
07			technology	16
08			interface	16
09			wiring	16
10			fieldbus technology	13

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 0, 0.33)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.047 \end{aligned}$$

Database and Middleware: Database and Middleware Introduction				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	middleware	0	database	24
02	database	24	data	18
03	integrity	10	relation	17
04	security	3	transaction	11
05			integrity	10
06			table	9
07			attribute	8
08			structure	7
09			unroll	5
10			property	5

Appendix D: Experimental Data

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0.25, 0, 0, 0.25, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.577 \end{aligned}$$

Database and Middleware: Entity-Relationship models				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	Entity-Relationship models	2	model	36
02	normalization	4	relation	17
03	amoralities	1	entity	14
04	relation	17	attribute	12
05			normalization	4
06			database	4
07			Entity-Relationship models	2
08			amoralities	1
09			object	1
10			cardinalities	1

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.25, 0, 0, 0.25, 0, 0.25, 0.25, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.795 \end{aligned}$$

Knowledge Engineering

Knowledge Engineering Overview I				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	logic	35	knowledge	97
02	predicate calculus	9	engineering	77
03	knowledge based system	9	system	47
04	knowledge representation	30	logic	35
05	future trends	0	knowledge representation	30
06			predicate	27
07			object	26
08			inference	21
09			manufacturing	18
10			data	17

$$sim(d, q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0.2, 0.2, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}}$$

$$= 0.346$$

Knowledge Engineering Overview II				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	rule based	39	object	108
02	knowledge representation	19	rule	84
03	reasoning	21	knowledge	56
04	uncertainty	24	representation	49
05	neutral network	17	rule based	39
06			class	32
07			hypothesis	25
08			uncertainty	24
09			data	22
10			reasoning	21

$$sim(d, q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0.2, 0, 0, 0.2, 0, 0, 0.2)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}}$$

$$= 0.289$$

Agent Introduction I: Why, When, and Where to Use Software Agents				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	agent	252	agent	252
02	software	58	software	58
03	agency	5	java	44
04	agent application	5	development	34
05	toolkit	13	intelligent	28

Appendix D: Experimental Data

06			environment	18
07			agentbuilder	16
08			architecture	16
09			distributed	16
10			toolkit	13

$$\begin{aligned} \text{sim}(d, q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0.2, 0.2, 0.0, 0.0, 0.0, 0.0, 0.2)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.577 \end{aligned}$$

Agent Introduction II: Agent-Based Systems for Intelligent Manufacturing: A State-of-the-Art Survey				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	agent	432	agent	432
02	survey	10	system	265
03	intelligent manufacturing	31	scheduling	172
04	agent based system	19	control	96
05	research	39	architecture	89
06			intelligent	71
07			planning	64
08			design	59
09			multi	45
10			chain	45

$$\begin{aligned} \text{sim}(d, q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0.2, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.289 \end{aligned}$$

Concurrent Knowledge Processing Techniques I				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	concurrent engineering	11	blackboard	114
02	blackboard	114	control	114
03	control blackboard	28	knowledge	86
04	distributed blackboard	1	problem	85
05	application	9	domain	40
06			scheduling	38
07			control blackboard	28
08			framework	26
09			cycle	26
10			mechanism	26

$$\begin{aligned} \text{sim}(d, q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0.2, 0.0, 0.0, 0.0, 0.2, 0.0, 0.0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.404 \end{aligned}$$

Concurrent Knowledge Processing Techniques II				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	distributed blackboard	9	blackboard	81
02	knowledge source	14	knowledge	76
03	multiprocessing	13	design	48
04	KSI	38	node	43
05	synchronizing	22	engineering	40
06			concurrent	40
07			system	40
08			KSI	38
09			distributed	24
10			synchronizing	22

$$sim(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 2, 0, 2)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}}$$

$$= 0.116$$

Autonomous Agents/Actors I				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	agent technology	5	agent	349
02	agent type	10	engineering	53
03	structure	9	knowledge	50
04	CIM system	1	reactive	33
05	multi-agent	4	collaborative	32
06			mobile	27
07			architecture	27
08			environment	27
09			interface	21
10			utility	18

$$sim(d,q) = \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}}$$

$$= 0.000$$

Autonomous Agents/Actors II				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	multi-agent	15	agent	204
02	CIM	18	task	153
03	Categorization of MAS	3	system	77
04	control type	0	knowledge	64
05	communication structure	3	control	47
06			communication	36
07			allocation	35
08			data	32

Appendix D: Experimental Data

09			scheduling	31
10			model	27

$$\begin{aligned} \text{sim}(d, q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.000 \end{aligned}$$

Machine Learning				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	inductive learning	8	learning	244
02	deductive learning	5	machine	144
03	analogical learning	1	attribute	72
04	neutral network	3	tree	65
05	genetic algorithm	3	knowledge	55
06	EBL	17	concept	48
07	CBR	16	decision	36
08			information	33
09			theory	32
10			entropy	27

$$\begin{aligned} \text{sim}(d, q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.000 \end{aligned}$$

Evolutionary Computation				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	genetic algorithm	10	fitness	47
02	optimisation	15	crossover	30
03	encoding	10	selection	22
04	mutation	8	genetic	20
05	crossover	30	generation	19
06			mating	17
07			chromosome	17
08			population	17
09			optimisation	15
10			offspring	12

$$\begin{aligned} \text{sim}(d, q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.2, 0, 0, 0, 0, 0, 0, 0.2, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.317 \end{aligned}$$

Data Mining and Knowledge Discovery				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	data mining	20	data	96
02	knowledge discovery	3	information	24

Appendix D: Experimental Data

03	task	7	clustering	23
04	estimation	7	analysis	23
05	prediction	12	data mining	20
06	clustering	23	object	14
07			model	14
08			pattern	12
09			knowledge	9
10			classification	9

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0.167, 0, 0.167, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.336 \end{aligned}$$

Autonomous agent - AI Planning				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	plan	90	action	151
02	autonomous agent	1	plan	90
03	IPEM	6	behaviour	68
04	reactive planning	1	precondition	42
05	hybrid architecture	4	execution	42
06	deliberative planning	2	unsupported	29
07			knowledge	25
08			robot	34
09			reactive	23
10			state	20

$$\begin{aligned} \text{sim}(d,q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0.167, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\ &= 0.217 \end{aligned}$$

Sub-Symbolic Learning				
No.	Concept Term	Frequency	Most Frequent Document Terms	Frequency
01	Kohonen Map	8	weights	124
02	Neural network learning	1	vector	112
03	pattern association	3	learning	77
04	Hebb Net	5	pattern	72
05	Perceptron	16	training	64
06	Hetero-Associative Memory (HAM) Network	1	layer	60
07	Backpropagation	10	cluster	59
08			network	51
09			node	32
10			matrix	26

Appendix D: Experimental Data

$$\begin{aligned}
 \text{sim}(d, q) &= \frac{(0.182, 0.164, 0.145, 0.127, 0.109, 0.091, 0.073, 0.055, 0.036, 0.018) \times (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)}{\sqrt{0.182^2 + 0.164^2 + 0.145^2 + 0.127^2 + 0.109^2 + 0.091^2 + 0.073^2 + 0.055^2 + 0.036^2 + 0.018^2} \times \sqrt{0.25^2 + 0.25^2}} \\
 &= 0.000
 \end{aligned}$$