International Journal of Pure and Applied Mathematics
Volume 12 No. 1 2004, 47-57

MATHEMATICAL/LOGICAL STRUCTURE OF PEDAGOGICAL TEXTS

M.T. Mahmoudi[1 §], K. Badie[2], N. Reyhani[3], J.A. Moghaddam[4]
[1,2,3,4]Info Society Department
Iran Telecom Research Center
End of North Karegar, Tehran, IRAN
[1]e-mail: mahmodi@itrc.ac.ir

Abstract: It has been demonstrated that educational texts can play a high role in achieving pedagogical purposes for both specialists and researchers. In this view, mathematical/logical considerations are of particular significance in organizing pedagogical texts. The related aspects which are addressed in this paper, include (i) mathematical/logical structure for conceptualization of the text features, (ii) mapping formalisms to justify/rationalize the relations between text features and learner model/style components on one side and abilities to be generated in the learner on the other side, and (iii) the mathematical/logical formalisms which are essential to assessing a text systematically through the reinforcement feedbacks from the learners.

AMS Subject Classification: 26A33
Key Words: mathematical/logical, pedagogical text, justification/realization, mapping formalism

1. Introduction

It has been demonstrated that text can play a promising role in achieving pedagogical purposes for a wide range of users such as simple learners, spe-

Received: January 10, 2004

§Correspondence author

cialists and researchers. In this view, mathematical/logical considerations are of particular significance in organizing pedagogical texts. These considerations can be related to (i) mathematical logical structures for conceptualization of the text features which can be viewed as an issue with respect to text representation, (ii) mathematical/logical structures rationalize the relations between text features and learner model/style on one side and abilities to be generated in the learner on the other side, and (iii) mathematical/logical structures, which are essential to assessing a text systematically through the learner's feedbacks so as to provide text revisions in an efficient way. In this paper, while emphasizing on (ii) and (iii), we elaborate the related mathematical/logical structures via raising the very formats of computation, which are essential to realizing issues such as mapping and assessment. The organization of the paper is as follows: Section 2 and Section 3 respectively address mathematical aspects for mapping, and mathematical formalism for assessment and revision of the pedagogical text. Concluding remark and future prospect has also come in Section 4.

2. Mathematical Aspects for Mapping and Assessment Issues in Pedagogical Texts

2.1. Types and Layers of Features Considered in Organizing the Pedagogical Texts

An ontology of features considered in organizing the pedagogical texts is illustrated in Figure 1. As it is seen from this ontology, two types of features: "over-all" and "detailed", are considered in organizing the text. With respect to "overall" type, titles of the sections, which are named compound issues, are constituted on the basis of coupling the scientifically significant concepts stated in terms of nouns or adjectives (modes), with the focal issue of the text, which is stated in terms of a verb. With respect to "detailed" type, following layers are considered: (a) introduction layer, (b) to-issue layer, (c) application layer, (d) advantages and disadvantages layer, (e) topic generation layer, and (f) reference layer.

2.2. Mapping from Researcher Space onto Text Space

The prime function of pedagogical text, as discussed earlier, is to provide suitable conditions for a spectrum of different learners with different approaches to

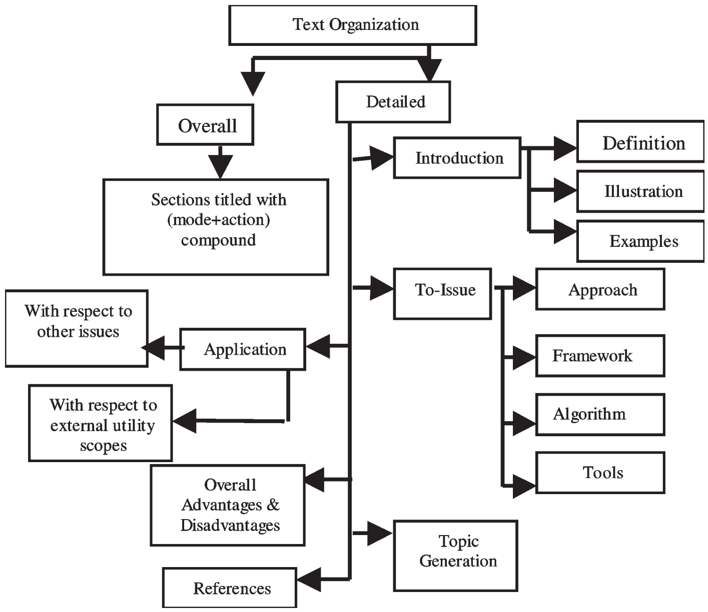

Figure 1: An ontology for the pedagogical text features

research in such a manner that certain abilities can be formed in them. This necessitates a systematic method for organizing the pedagogical text from the viewpoint of its chapters/sections, and the features, which are to be considered in organizing these chapters/sections. Researcher's approach to research, in its simplest form, can be described in terms of the types of research he/she is intending to conduct. Types of research in turn call for a variety of reasoning styles to be fulfilled. Reasoning styles can therefore represent approaches to research in some way. Figure 2 illustrates the relations between types of research, and the reasoning styles considered in our study. Also, types of abilities in research, considered in our approach to organizing pedagogical texts, are shown in Figure 3. With respect to mathematical/logical structures for rationalization of the relations between text features and learner model/style on one side and abilities in the learner on the other side, it is important to know how the ensemble of a certain research ability and a certain reasoning approach can be projected onto a number of considerations with respect to text features at different layers of the illustrated ontology. It is obvious that, knowing this projection from the beginning is almost impossible, since there is no domain theory, which can support the fact in an analytical way. One instead should rely on the fact that there are a number of patterns, each describing the associations between the ensemble of a certain research ability and a certain reasoning approach as a situation, and some considerations in text features as the related solution. According to this view, new associations can be constructed on the basis of the known associations using a method, which has the capability to do interpolative reasoning. Since the number of known associations is not high, an intelligent method with learning capability is preferred. Case-Based Reasoning can be a good alternative in this regard [6]. Here, the case library can be built using either experienced association patterns, or those patterns which seem to be sound from the expert/specialist's viewpoint. Regarding the above discussion, when a researcher with certain reasoning style(s) is faced with a situation within which certain research ability is demanded, similar cases existing in the library are first retrieved, and their solutions, which are the corresponding considerations in text features, are then composed in a way that the resultant can hold minimum distance toward the expected text features for the current situation. It is in this way that the final considerations respecting the text features will be shaped.

Figure 4 (a) illustrates a current situation in research together with some examples for the retrieved cases. The similarity assessment function to be used for case retrieval can be a simple Euclidean distance function, [3]. To perform compositional adaptation, a method based on using global distance function [8] can be utilized. According to this method, this approach tries to compose the retrieved cases in order to find an appropriate solution for ongoing problem through optimizing the objective function (1). The basic idea behind GDF is that this similarity between the ongoing problem and the retrieved cases should be in a small vicinity of the similarity of the final solution and retrieved case solution. Based upon the above, the following object function is defined.

$$\sum_{i=0}^{l} \left(\mathrm{sim}\left(\mathrm{sit}^{\mathrm{cur}}, \mathrm{sit}^{i}\right) - \mathrm{sim}(\mathrm{sol}^{\mathrm{cur}}, \mathrm{sol}^{i})\right)^2 \tag{1}$$

Where l is the number of retrieved cases, $\mathrm{sit}^{\mathrm{cur}}$ is the current problem, sit^{i} is the situation of i-th retrieved case, $\mathrm{sol}^{\mathrm{cur}}$ is the solution of the ongoing problem, and sim is the similarity function which is used during case retrieval phase. Having applied the global distance function, the final solution for the current situation of Figure 4 (a), based on composing the solutions of the retrieved cases, is shown in Figure 4 (b). It should be noted that the solutions derived the above manner may not necessarily be sound due to a variety of reasons such as the very bias existing in the formats of case representation, retrieval and adaptation. Here, based upon the feedbacks from the researcher as the user

Research types / Reasoning style	Comparative modeling	Theoretical analysis	Scenario making	Prospect anticipation	Text organization	Product customization	Framework proposition
Formal reasoning	✓	✓					✓
Domain theory based reasoning	✓	✓	✓	✓	✓	✓	✓
Experimentation					✓	✓	✓
Scenario-based reasoning			✓	✓	✓	✓	✓

Figure 2: Relationships between types of research and the reasoning styles

Referencing
selection
interpretation
reconsideration
topic generation
valuation & verification

Figure 3: Types of research abilities

of the text, some kind of learning from experience is to be performed, to assure a lower error rate for the next derivations of text features. Regarding this, a strategy making use of a secondary case-based reasoning for adjusting the bias values of the primary modules of retrieval and adaptation, can be used [2], [9], [1].

3. Mathemnatical Formalism for Assessment and Revision of the Pedagogical Text

3.1. Mathematical Formalism for Assessing the Pedagogical Text

It is obvious that, the initial cases may not necessarily be valid or exact due to a variety of reasons such as expert's lack of knowledge with respect to research situations or feature requirements for particular types of research, the uncertainty which is unavoidable in describing some of the cases, and occasional incorrectness of his/her statements about the way a text is to be utilized by its users. Due to this fact, it is occasionally necessary to revise the way feature derivation is to be performed. It was discussed in Section 2 that a secondary case-based reasoning can gradually lead text organization to the stage, where the final text features can become reasonably effective for the researcher, due to the point that the bias values for the primary modules parameters can be adjusted well. Another perspective based upon case optimization could however exist to approach revision systematically, using Genetic Algorithms [4]. Regarding this approach, each set of features for a certain research ability is regarded as a chromosome in which the information of each feature constitutes a gene. Based upon the feedback describing the effectiveness of the exposed pattern of features, a fitness value is calculated. Here, the fitness function can be a simple error function whose variables are related to the system output and desired one. Using the well-known operators of selection, cross-over and muta-

Situation (Retrieved Case #1); Distance=.645	
Validation: L Reconsideration: L Selection: M Interpretation: H	Research Ability
Formal Reasoning: L Qualitative Domain Theory-Based Reasoning: H Scenario Based Reasoning:H Experimentation Reasoning: L	Reasoning Style
Solution	
Approach: M Framework: L Algorithm: H Tools: L	

Situation (Retrieved Case #2) ; Distance=.355	
Validation: L Reconsideration: L Selection: H Interpretation: H	Research Ability
Formal Reasoning: L Qualitative Domain Theory-Based Reasoning: M Scenario Based Reasoning:H Experimentation Reasoning: L	Reasoning Style
Solution	
Approach: H Framework: L Algorithm: H Tools: L	

Ongoing Problem	
Validation: L Reconsideration: L Selection: H Interpretation: H	Research Ability
Formal Reasoning: L Qualitative Domain Theory-Based Reasoning: M Scenario Based Reasoning:M Experimentation Reasoning: L	Reasoning Style

Figure 4: (a) Current situation in research and some examples for the retrieved cases

Solution
Approach: M Framework: M Algorithm: H Tools: L

Figure 4: (b) Find solution to be obtained through composition

tion, some chromosomes are finally obtained with reasonable amount of fitness values showing which certain features, are to be considered for a researcher who intends to acquire a certain research ability for a certain approach to research. A salient difference between the two approaches discussed above, is that genetic algorithm presupposes ample of sample problems through which optimal structures with respect to the features can emerge for different combinations of reasoning approach and research ability. However, in case of using CBR approach, it is presupposed that, due to some environmental limitations, having a high number of cases as sample problems is not feasible. Regarding the above point, text revision can ultimately be based on a policy as follows: initial cases are constituted for different research situations, and feature derivation process is then performed for new situations, based on these cases and process of compositional adaptation. In the meantime, the entire process is controlled by a secondary case-based reasoning, to adjust the parameters values for case retrieval and adaptation. If it was observed that the total trend of derivation is getting worse, which means that the features derived in this way are not totally satisfactory for the researchers, and at the same time ample of cases have already been obtained, it would then be natural to think that a holistic optimization approach is necessary to search for the optimal structures of features for different research situations whose related cases already exist in the case library. At this stage, genetic algorithm is activated to derive these optimal structures.

It is therefore noticed that, from a mathematical point of view, CBR and Genetic Algorithm act complementarily to each other to ensure a total compromise between the cost of providing sample problems on the one hand, and the cost for optimization on the other hand. This in general can be regarded as a helpful guide-line for many interpolation problems.

3.2. Reducing the Cost of the Feedback to be Received from Research for Text Revision

It was discussed that Genetic Algorithm (GA) can be a suitable choice for optimizing the feature patterns, which are to be exposed to researcher for certain combinations of reasoning style and research ability. It should however be noticed that each feedback from researcher needs a minimum amount of cost that itself is a negative impact on feasibility of applying GA as a practical method for optimizing the feature patterns. To circumvent this problem, one may think of the possibility of reducing the cost of the feedbacks through estimating the fitness of chromosomes by an estimator which has already been trained on the basis of a variety of genetically produced chromosomes for which real feedbacks have been given by researcher. Regarding this, optimization process starts from randomly selecting some chromosomes and calculating their fitness value based on the feedbacks from the researcher. At this stage, neural network and preferably a radial basis function neural net (RBFNN) may be trained based on the related information, to take the responsibility of estimating the fitness values for the chromosomes to be generated later. The main reason for selecting RBF as a ground for the estimator, is its capability in clustering the input chromosomes based on their status of belonging to the same normal distribution.

Once the estimator was built by the above manner, Genetic Algorithm would them be applied to produce new chromosomes based on the previous ones. The fitness values of newly produced chromosomes will then be estimated using the above estimator, and the process of applying GA will again be repeated. Here, to get assured of the functional reliability of the estimator, some of its intermediate outputs are compared with the real values extracted from the feedbacks, and based on the related error, the estimator will be readjusted. Readjustment process is performed in the following manner: Error variance is calculated first, and in case that the value is higher than a certain threshold, a test called -test will be activated to retrain the estimator in such a manner that this variance can get normal. The process of applying -test is in reality based on a sort of selection out of the existing training patterns whose error variance rate gets sufficiently small [11], [10]. This process is continued until the stage, where the estimator is found to be capable of estimating the fitness values with a tolerable amount of error. At this stage, the estimator is almost ready to estimate the chromosomes fitness values with a high confidence level. Therefore, it will function until the stage where some chromosomes are found whose fitness values are high than a certain confidence level.

4. Concluding Remarks

In this paper, we elaborated some essential mathematical/logical structures for organizing pedagogical texts, emphasizing on mapping from researcher space on to texts space as the kernel idea. In this respect, after reviewing some existing approaches to pedagogical texts from mathematical/logical points of view, mathematical structures for mapping from researcher space onto text space as well as text assessment and revision were discussed. Regarding this, case-based reasoning was mentioned to take the responsibility of deriving the feature pattern for a certain situation in researcher using a sophisticated compositional adaptation algorithm based on global distance function. This will eventually lead to realization of the projection from researcher space onto text space. Although the secondary CBR has the ability to adjust the bias values of the global distance function, and thus assuring the functional robustness of the primary case-based reasoning module, possibilities however exist that CBR by itself can not take the responsibility mentioned above. Due to this reason, genetic algorithm was proposed as a complementary means for optimizing the final formats of projection from researcher space onto text space. To reduce the cost of feedbacks necessary for implementing the genetic algorithm, the concept of error estimation for estimating fitness values using radial basis function neural net was employed. It is seen that, the collaborative employment of CBR, GA and RBFNN can lead to an environment in which both feature pattern derivation (for decision making) and assessment (for learning) are being performed in a conjunctive manner, thus becoming practical for real world problems like pedagogy.

It is seen that, the mathematical structure suggested in this paper for organizing pedagogical texts, can be equally applied for pedagogical purposes in any domain. Also, due to its domain-independence, it can be used successfully for any types of service environment, in which some packages of service are to be designed to serve the end users with different mental models.

References

[1] A. Ahmadi, A.K. Lamjiri, M. Nevisi, K. Badie, Using a two-layered case-based reasoning for prediction in soccer coach, In: *Proc. of MLMTA'2003.*

[2] K. Badie, N. Reyhani, Learning through estimating optimal formats for problem solving modules, In: *Proc. of IEEE International Conf. on Systems*, Man. and Cybernetics, Tunisia (2002).

[3] B.V. Dasarathy, *Nearest Neighbor Norms: NN Pattern Classification Techniques*, IEEE Computer Science Press (1990).

[4] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (1992).

[5] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Second Edition, Prentice Hall (1999).

[6] M. Lenz, Case Based Reasoning Technology: From Foundation to Applications, Springer-Verlag (2002).

[7] M. Norgaard, O. Ravn, N.K. Poulsen, L.K. Hansen, *Neural Network for Modeling and Control of Dynamic Systems*, Springer-Verlag (2000).

[8] N. Reyhani, K. Badie, M. Kharrat, A new approach to compositional adaptation based on optimizing the global distance function and application in an intelligent tutoring system, In: *IEEE International Conf. on Information Integration and Reuse* (2003).

[9] N. Reyhani, K. Badie, M. Kharrat, A two layered case-based reasoning approach to text summarization pattern, In: *Proc. of IEEE Systems and Information Engineering Design Symposium* (2003), 47-51.

[10] N. Reyhani, K. Badie, M. Kharrat, A new approach to heterogeneous time series analysis using hybrid case-based reasoning and additive fuzzy systems, *WSEAS Transaction in Systems*, **2**, No. 4 (2003), 991-996.

[11] A. Tsui, A. Jones, A. Oliveria, The construction of smooth models using irregular embeddings determined by a gamma test analysis, *Neural Computing and Applications*, **10**, No. 4 (2002), 318-329.