Chapter 9: Summary and Speculations

In this dissertation, a new theoretical and mathematical framework is developed in attempt to build a coherent model to study the student learning of physics. The theoretical foundation is based on the many ideas and theories developed by researchers in cognitive science and physics education. In the following sections, I summarize several key elements in the new theory with emphasis on the numerical methods. I will not discuss the applications in the study of student understanding of quantum mechanics.

Key Elements in the New Theory

Student Models

In our research, we describe the different student views and understandings of physics concepts with "student models". A student model is defined as a productive mental structure that can be applied to a variety of different physical contexts to generate explanatory results. The formation as well as the application of a model has strong involvement of physical contexts, i.e., the models are highly context dependent.

Physics Models

For a particular physics concept, through systematic research, we can identify a finite set of commonly recognized student models. These models usually consist of one correct expert model and a few incorrect or partially correct student models. These models are defined as *Physical Models* since they are common to a group of students with similar background and the existence of these models can be verified repeatedly through research.

Student Model State

For a single student solving a set of problems related to a single physics concept, there are usually two different situations:

1. The student can use one of the physical models and be consistent in using it in solving all questions. The model can be either the expert model or another physical model (e.g. an incorrect student model).

2. The student can hold different physical models at the same time and be inconsistent in using them, i.e. the student can use one of the physical models on certain questions and use another one on other questions although all the questions are related to a same concept and the questions are seen as equivalent by experts.

Then the different situations of students using their models are described with different student model states. The first case corresponds to a *consistent model state* and the second case is considered as a *mixed model state*.

With a set of questions designed around a single physics concept domain, we can measure the probability for a single student to be triggered into the different physical models in solving these problems. For different students, the distributions of probability will be different. Therefore, we can use these probabilities to represent student model state. Thus, the student model state is defined *as a specific configuration of probabilities for a student to use different physical models in problem-solving contexts related to a particular physical concept domain*.

Random Process in Model Triggering

How a student is triggered into a specific physical model is a very complicated process. It depends on both the students' background and the structural information of the questions. The situation can be even worse when we study a large population of students with diverse backgrounds. In this dissertation, I will not go into the details of model triggering but rather characterize it as a conditioned random process that is constrained by the background of the students and the physical features of the questions.

The meaning of this conditioned random process is defined as the following: *In a well-defined domain of physics context related to one physics concept, students can come up with a finite set of models to deal with the various problem solving situations within the context domain. What type of model is to be triggered by a specific physical context is a random process to an external observer, however, the set of the possible models is bounded and known (these are defined as physical models).*

Model Space

For a specific physics concept domain, we can represent the complete set of physics models with a set of orthogonal unitary vectors, $\mathbf{e}_\eta$, defined as the physical model vectors:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} ; \ \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} ; \ \ldots \ \mathbf{e}_w = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \qquad \textit{(2-2)}$$

where "w" is the total number of physical models associated with the specific concept under consideration (a null model is also included).

The space spanned by all these physical model vectors is defined as the *model space*. Then the model state for a single student, the k^{th} student in a class, can be represented with a unitary vector $\mathbf{u}_k$:

$$\left| \mathbf{u}_k \right\rangle = \begin{pmatrix} \sqrt{q_1^k} \\ \sqrt{q_2^k} \\ \vdots \\ \sqrt{q_w^k} \end{pmatrix} \qquad \textit{(2-3)}$$

where $\langle \mathbf{u}_k | \mathbf{u}_k \rangle = \sum_{\eta=1}^{w} q_\eta^k = 1$ *(2-4)*

This normalization is constrained by the requirement that the physical models form a complete set and the probability for a student to be triggered into a model in the set is therefore 1.

Model Analysis

Model analysis consists a set of numerical algorithms developed based on the theory. The algorithms are designed to provide quantitative evaluations of student models with research-based multiple-choice questions. In the following sections, I specifically summarize two basic algorithms. In chapter 5, additional methods are discussed to deal with specific features of different types of multiple-choice interments.

- **Model Evaluation**

With a set of research-based questions, student responses can be used to infer the models they used to solve the questions. Using a certain set of questions associated with a single physics concept domain (e.g. Newton III), we can obtain the single student model state as:

$$\mathbf{u}_k = \begin{pmatrix} u_{1k} \\ u_{2k} \\ u_{3k} \end{pmatrix} = \frac{1}{\sqrt{m}} \begin{pmatrix} \sqrt{n_{1k}} \\ \sqrt{n_{2k}} \\ \sqrt{n_{3k}} \end{pmatrix} = | \mathbf{u}_k \rangle \qquad (4\text{-}3)$$

where k represents the kth student and *m* is the number of questions used in the measurement. In the equation, $n_{\eta k}$ respondents the number of questions answered with physical model η by the kth student. Then we can construct the single student model density matrix of the kth student as:

$$\mathcal{D}_k = | \mathbf{u}_k \rangle \langle \mathbf{u}_k | = \left\{ \rho_{\eta\mu}^k \right\} = \frac{1}{m} \begin{bmatrix} n_{1k} & \sqrt{n_{1k} n_{2k}} & \sqrt{n_{1k} n_{3k}} \\ \sqrt{n_{2k} n_{1k}} & n_{2k} & \sqrt{n_{2k} n_{3k}} \\ \sqrt{n_{3k} n_{1k}} & \sqrt{n_{3k} n_{2k}} & n_{3k} \end{bmatrix} \qquad (4\text{-}4)$$

which can be used to form the class model density matrix.

$$\mathcal{D} == \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{33} \end{bmatrix} = \frac{1}{N} \sum_{k=1}^{N} \mathcal{D}_k = \frac{1}{N} \sum_{k=1}^{N} \begin{bmatrix} \rho_{11}^k & \rho_{12}^k & \rho_{13}^k \\ \rho_{21}^k & \rho_{22}^k & \rho_{23}^k \\ \rho_{31}^k & \rho_{32}^k & \rho_{33}^k \end{bmatrix} \qquad (4\text{-}6)$$

By analyzing the class model density matrix, we can study the features of the models used by the students in the class. A graphical representation, the model plot, is often used to present the results of student models (See figure 9-1).

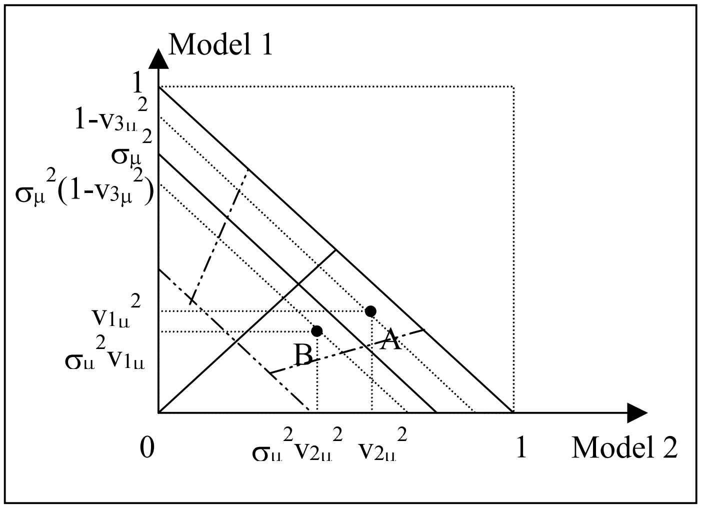

Figure 9-1. Example from chapter 4 (figure 4-6)

- **The Concentration Factor**

Based on our understanding of student learning, the student responses are considered as the output of students applying their models in various physical contexts. Therefore, if students have a few consistent models of physics, the responses should be more concentrated on those choices representing the corresponding models. On the other hand, if students have no models, or have a wide variety of models and use them inconsistently, their responses will be more randomly distributed among the choices. Therefore, the way in which the students' responses on multiple-choice questions are distributed reveals important information on student models.

To study the student responses, it is convenient to have a simple measure that gives the information on the distribution of the responses. It is defined as the concentration factor, denoted as ***C***, with a value in the range [0,1], where a large value represents highly concentrated responses.

To formulate ***C***, first sum up all the student responses on one question in a vector form and get the total response vector for that question:

$$\vec{r} = \sum_{k=1}^{N} \vec{r}_k = (n_1, n_2, \ldots, n_i, \ldots, n_m)$$

where n_i is the total number of students who selected choice i. Since there is a total of N responses, we have

$$\sum_{i}^{m} n_i = N \qquad (3\text{-}1)$$

Define the scaled length of $\vec{r}$ as

$$r_0 = \frac{\sqrt{\sum_{i=1}^{m} n_i^2}}{N}$$

where

$$\frac{1}{\sqrt{m}} \le r_0 \le 1$$

This suggests to choose ***C*** as:

$$\boldsymbol{C} = \frac{\sqrt{m}}{\sqrt{m}-1} \times (r_0 - \frac{1}{\sqrt{m}}) = \frac{\sqrt{m}}{(\sqrt{m}-1)} \times (\frac{\sqrt{\sum_{i=1}^{m} n_i^2}}{N} - \frac{1}{\sqrt{m}}) \qquad \textit{(3-2a)}$$

where N >> m is required.

To use the concentration factor, the first method is to combine the concentration factor with scores to form response patterns. A low score and low concentration type shows that most of the students have no dominating model on the topic (at least as revealed by the test being studied) and their responses are more like the results of random guesses. On the other hand, a low score but high concentration type implies that the students probably have a strong incorrect model on the related concept. If the results are from a pre-test, the instructors can be informed by these incorrect initial student models and prepare for appropriate instruction.

We can also represent the results on a "S−***C***" plot, using the score as the horizontal axis and the concentration as the vertical axis. Then the response of each question could be represented as a point on the S−***C*** plot. The shift of the response can be represented with a vector starting from the point representing the initial state towards the point for the final state.

Application Examples

To illustrate how the algorithms are applied, I summarize two examples with FCI data from UMd students. These two examples are all on the concept of Force and Motion relation, which involve 5 FCI questions are associated with this concept (questions 5, 9, 18, 22 and 28).

Student Class Model States

As discussed in chapter 2, student responses to the five FCI questions involve three physical models:

Model 1: It is not necessary to have a force to maintain motion and there is no such thing as a "force in the direction of motion". (Correct)

Model 2: A force is needed to maintain motion. This model also includes the ideas that there is always a force in the direction of motion and that the force is directly related to the velocity of motion. (Incorrect)

Model 3: Other ideas and incomplete answers. (A null model)

For the three physical models in FM group, our associations of the responses corresponding to the FCI questions in FM cluster are listed in table 9-1.

Table 9-1. Modeling of student responses for FM group (table 4-1)

Question	Model 1	Model 2	Model 3
5	d	a, b, c	e
9	a, d	b, c	e
18	b	a, e	c, d
22	a, d	b, c, e	
28	c	a, d, e	b

Following the algorithm procedures, we can construct student model density matrix and perform eigenvalue decomposition on it. The primary student class model states are shown in figure 9-2. The data is from 7 tutorial classes and 7 traditional classed in UMd.

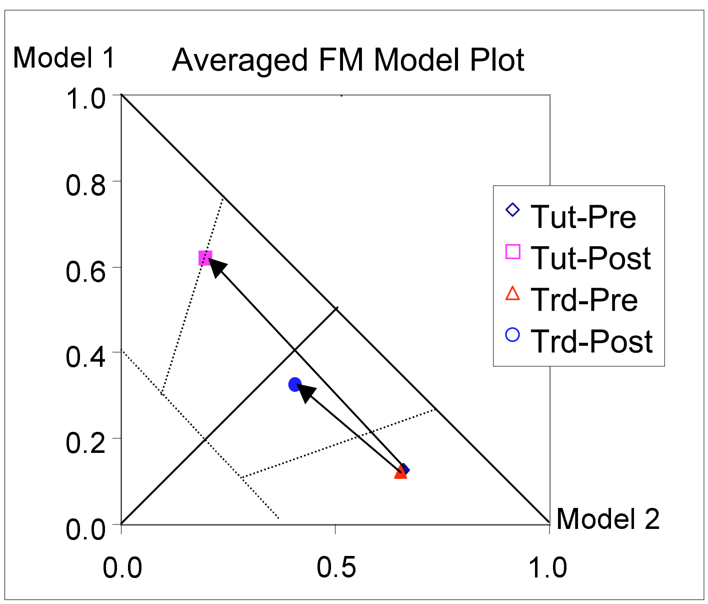

Figure 9-2. Student models on Force-Motion (FCI, UMd, figure 4-10)

As we can see the initial states of both types of classes fall into the model 2 region, which indicates that the students are having a consistent but incorrect model (strong misconception) – “there is always a force in the direction of motion”. After the instruction, the tutorial classes show some promising improvement towards the favorable model. Although still in the mixed model region, the tutorial classes make a quite large favorable shift comparing to its problematic initial state. On the other hand, the situation for the traditional classes is not so optimistic. Its final state doesn't even cross the centerline, i.e., the primary model fails to make the transition to the favorable side. This situation indicates that many students are still in favor of their initial misconception. Since the final state is also very close to the centerline, the student model can be interpreted as very mixed and inconsistent under equal influences from the correct physical model and the initial misconception.

Concentration of Student Responses on FCI

The pre-instruction FCI data of 16 UMd classes are analyzed with a three-level modeling scheme described in chapter 3. The results are very similar for all classes. This is expected since the background of the incoming students is similar. Therefore, the results of the pre-data analysis are combined. Table 9-3 is a list of the coding of the pre-test response types for all 29 questions on the FCI test.

Table 9-2. Pre-instruction FCI response types (UMd, table 3-5b)

Types	LL	LM	LH	ML	MM	MH	HH
Patterns	No peak	Two peaks	One peak	Weak one-peak	Two peaks	Weak one-peak	One peak
Questions	15, 24	5, 9, 18, 28	2, 13, 22	3, 7, 21, 26	6, 8, 11, 14, 17, 20, 23, 25	12, 16, 29	1, 4, 10, 19, 27

A look at the details of the questions suggests that most of the questions with LM and LH types are dealing with two physics concepts, the relation of force and motion and Newton III (see chapter 2 for details on the two concept groups).

Advantages of Model Analysis

This method makes more use of the data than score-based analysis, (traditionally if we only calculate the correct answers, a lot of information is wasted) and allows us to study the student models in a quantitative way. It can serve as a more comprehensive quantitative evaluation for the student performance, especially with respect to their mental models.

With this method, student raw data are transferred into states in the model spaces. Since for each physics concept domain, the structure of the associated model space is definite, the student model states obtained with different instruments can be directly compared in the model space. The results can be used to analyze student understanding and/or the features of the instruments. Model analysis can also provide more explicit information on how to improve instruction. Since it gives more detailed knowledge of the models the students have, it allows us to see more directly about the possible causes of the student difficulties rather than just the difficulty itself. Therefore we can develop more appropriate instruction strategies right to the weak points and help the student more effectively.

Speculations on Modeling the Learning Process

What tool can we use to effectively model the learning process? Many instruments and methods have been developed for various contexts. A non-conclusive list includes open-ended conceptual questions, student interviews, multiple-choice surveys and concept mapping.

A Crucial Problem on Precise Modeling of Student Understanding

All these methods are influenced by an important factor in the measurement, the communication between the students and the investigators (or the language issue for simplicity). The student actual thinking can be miscommunicated, i.e., the students often use the "words" or physical terms in ways different than what experts would expect. Students and experts can use the same language but mean completely different things. Therefore, student responses coded with the representations of simple language or other similar formats (e.g. links in concept maps) can cause problems in convey the actual thoughts used to generate those responses.

The Available Methods

- **Student Interviews**

A well-conducted interview is by far the most effective method to identify the "true" reasoning of students. In interviews, the investigator can have the opportunity not only to collect what the student are saying but also to probe in real-time what the students do mean by using those words. Detailed analysis of a good interview can provide the most insightful information on the student understandings.

Although student interview is an effective method, it is very time consuming to conduct the research and to analyze the data. A well-distributed population of students for the interview is also difficult to obtain. In addition, the experience of the investigators and the quality of the interview protocols can also affect the results significantly. All these make it impractical to use interviews as a general evaluation tool in instruction.

- **Open-ended Conceptual Questions**

Open-ended conceptual questions, depending on the design of the questions, can provide comparatively rich information on student reasoning. However, it still bears the problem of miscommunication, since we don't have a second chance to further probe the students when we have confusions on what they are saying. Neither can we investigate if what they say really means what it is meant to be. The data analysis of the open-ended questions is also time-consuming, which makes it difficult to implement this method in large classes.

- **Multiple-choice Questions**

Multiple-choice questions often provide the least information on student thinking especially when using score based evaluations. If the questions are not properly designed, the results can even provide misleading interpretations. However, multiple-choice instrument does have many advantages. The data collection and analysis are easy to conduct which make it an ideal instrument for large classes.

- **Concept Mapping**

Concept map often raises a lot of disputes on what actually is being measured. As far as the "communication" issue is concerned, when students draw links between different concepts denoted by "words" on the paper, which the students have heard for many times, it is fairly difficult to decide what those links actually mean. For example, suppose we give students four concepts described with "waves", "amplitude", "frequency", and "wavelength". Through pure memorization, many students know that all the "words" are related (regardless what these words actually mean to the individual students). Therefore, the links in the concept maps constructed by the students can be quite similar to those from experts (experts and novices can have differently structured concept maps, but the information that can be extracted from such structures is limited and does not provide applicable guidance). What can this tell us about? It seems that a simple concept map only give the associations between different "words" memorized by the students. The links themselves does not provide much information, if any. What is important is the quality of the links, i.e., the content of the understandings that the links represent, and it is unlikely to represent such subtle issues with lines between words. Although by analyzing the structure of the whole concept map, we can obtain certain information on the student knowledge structure. Concept mapping often fail to provide precise information on specific student understandings.

The Method of Model Analysis

Model analysis integrates student interviews and multiple-choice questions together. The instrument, the model-based multiple-choice test, can be easily implemented in large classes to obtain quantitative evaluations on student models and the results are validated by research. The key elements of this method include the followings:

1. Through systematic research and detailed student interviews, common student models are identified and validated. The student models are reliable for a large population of students with similar background (e.g. the students in calculus-based introductory physics classes). These student models are then used to model the learning of students with similar background.

2. The multiple-choice questions used with model analysis are developed based on the knowledge of student models identified through qualitative research. The choices of the questions are designed to represent the common students models and are validated through research.

3. The numerical algorithms use the student full responses rather than just the scores. Student responses are analyzed in model space and the results can give explicit information on the student understandings (models).

Model analysis relies heavily on qualitative methods. By conducting systematic research, it is expected that the identified student models reflect the majority of different types of student understandings, and the multiple-choice instruments do not contain significant communication problems. In other words, we use interviews to identify the student actual thinking under certain contexts and use multiple-choice instruments to measure the students' using of these identified different understandings. The combination of the two methods can partially solve the communication problem and provide an effective and stable tool to probe large classes.

Once a reliable package is developed, instructors with some training can easily implement model analysis instruments in large classes to obtain immediate feedback from students with comparatively rich information on the student actual understandings.