# Piano Tutor: An Intelligent Keyboard Instruction System


Dr. Marta Sanchez
Dr. Annabelle Joseph
Dr. Roger Dannenberg
Dr. Philip L. Miller
Peter Capell
Robert Joseph


8 February 1988


*edited by*
Peter Capell
Robert Joseph


This is the technical portion of a proposal. An updated version intended for publication is in preparation.

## 1. Introduction

The Piano Tutor is a multi-media workstation that teaches basic piano playing. Playing the piano requires psychomotor[1] skills as well as intellectual understanding. Three emergent technologies make Piano Tutor an interesting and viable research opportunity including: recent developments in real-time music technology, expert tutoring systems, and the increased availability of high-capacity optical storage media. In addition, instructional systems design, a discipline for formal curriculum planning, training and development, is an integral part of the development of the system. This project gives us the opportunity to examine many interesting research topics and issues including:

- Music recognition software;
- User interfaces for teaching psychomotor skills;
- Intelligent Computer-Aided Instructional Systems for teaching psychomotor skills;
- Exploration of videodisc as a means of communication;
- Instructional systems design in relation to ICAI.

The Piano Tutor operates in two modes; as a teacher and as a silent observer/evaluator during practice. In its Teaching mode, Piano Tutor guides students through a dynamically changing lesson path with appropriate interruptions for suggestions and corrections. In its Practicing mode, Piano Tutor provides guidance without interruptions, providing students with graphic evaluations of their playing and among other capabilities.

Students may work with a teacher or by themselves, receiving expert piano instruction at a pace set by their own ability. Mastery of basic techniques becomes the student's private responsibility and a personal challenge.

This paper gives a brief overview of the process of learning to play the piano, description of the project in general outlining relevant issues and what is currently being done as well as a description of the unique opportunities this project to offers in a potentially rich learning environment.

## 2. Keyboard Instruction: Master and Apprentice

In contrast to most academic disciplines, music instruction is mainly conducted on a one-to-one basis in the tradition of master and apprentice. This is not surprising given the wide variety in student's music abilities in music and the blend of complex cognitive, physical and affective behaviors required for musical performance. Only in this very personal instructional setting can a teacher attend to the specific needs of each student. Our observations have shown that teachers provide a very "tailored" curriculum to each of their students:

- Pointing out errors and diagnosing their causes;
- Offering alternative solutions to technical and musical problems;
- Demonstrating concepts through examples;
- and selecting appropriate lessons according to the student's level of performance and taste.

---

[1]"Psychomotor" is a term from psychology, used to distinguish physical activities from cognitive ones. Operationally, solving math problems is cognitive, while playing the violin is considered to be psychomotor.

The role of the student is to remember and make use of explanations, demonstrations and suggestions given at the lesson and apply them in practice.

## 2.1. Objectives

Our most basic objectives are to combine a functional user interface, expert instructional diagnostics and music recognition software to provide students with expert piano instruction. This requires the system to be able to identify problems and select appropriate remedial strategies in a realistic way. The precision of Piano Tutor's judgements depends on the sensing of student performances with a combination of hardware and intelligent software. Through a combination of live-action sequences on videodisc and several hours of a teacher's voice, we hope to give Piano Tutor a high level of responsiveness, allowing it to realize a complex instructional plan.

The challenge for the Piano Tutor project is to emulate the precision of the expert piano teacher's diagnoses and to interact with the student effectively. Many specialized teaching methods cannot be a part of the Piano Tutor; such as actually moving the student's hands, observing the student's posture or identifying clumsy fingerings. Nonetheless, we *can* determine if the student is playing the correct note, the intensity with which each note is played, and the length of time it is held. These parameters serve as our means of diagnosing pitch errors, poor dynamics and rhythm problems. Beyond these basic elements, we have an obligation to characterize the teacher/pupil interaction in a reasonable way within the limits of technology. In the following sections we outline the software components and strategies we have employed in representing this interaction effectively.

## 2.2. Teaching the Piano

Typically, a piano lesson starts with the student playing a piece that has been practiced since the previous lesson. The teacher evaluates the performance and offers technical and musical suggestions. If subsequent performances are required, usually the teacher will stop the student to work on specific problems which can range from bad posture, awkward hand position, and playing wrong notes, to a lack of musical expression. Problems may be addressed by explanation, demonstration, simple reminders, use of alternate fingerings and specific suggestions for practice. Before the lesson ends, the teacher may play the new pieces assigned so that the student can hear how they should sound. The student's responsibility is regular practice with attention to suggestions, corrections and other information given in the lessons.

Practice is essential for learning to play the piano. Traditionally, the student receives no feedback during the practice period, and may not be aware of playing wrong notes or rhythm or having forgotten the instructors comments from the previous lesson. This slows the learning process and can be a source of frustration to both student and instructor. Efficient and effective practice is an integral part of the design of the Piano Tutor.

## 2.3. Multimedia Real-Time Interaction

We have an unprecedented opportunity to explore the possibilities of multimedia real-time interaction in computer-based instruction systems. Piano instruction requires a high degree of interaction both visually and aurally. This information includes video presentations of correct performance technique, music notation, text instructions, music and speech. All of these can take place in real-time; that is, the timing of the information is often as important as the content. An example would be the use of a spoken comment to interrupt a student when he or she makes a mistake.

Videodisc technology will be used to support an active and friendly interactive learning environment. A typical use of the videodisc will be to introduce new material at the beginning of a lesson. For example, a teacher might explain how to place the hand on the keyboard and then play a simple melody. The videodisc will also be used for remediation. For example, if a student is having difficulty playing with uniform loudness, the videodisc could illustrate the proper playing technique.

Unlike conventional interactive videodisc systems where computer programs are contained on the disc and run by special players, we intend to treat the videodisc player as a simple random access video playback unit. The intelligence will reside in a powerful computer to which the videodisc player is attached. We believe this configuration will greatly simplify our development. Section 4.1.5 describes how the videodisc fits into our software architecture. We will use the computer for digitized speech playback and we will have other means for synthesizing music. This means that we can revise spoken instructions and music interactively without the delay and expense of pressing a new disc.

The key to developing an educationally sound videodisc is that its structure be clearly defined in conceptual terms before production begins. For this reason, instructional design is critical to the design of a videodisc-based instructional system. We are fortunate to have an instructional designer (Peter Capell) who also has experience in all phases of video preproduction, including content and task analyses, scripting, and image design. Peter has additional experience as a producer.

The Carnegie Mellon University Software Engineering Institute has recently installed a video production center geared toward producing interactive videodiscs for software education and technology transfer. This new Betacam facility is available to us as a resource not only through regular University channels, but through our relationship to principle personnel associated with the studio, both within Carnegie Mellon and with the Software Engineering Institute. Additionally, the University Library's Instructional Technology Department has Betacam offline editing equipment which is available to us upon request.

## 3. The Piano Tutor

The work station we are building is composed of a computer connected to a velocity sensitive, piano-like MIDI[2] keyboard, MIDI synthesizer and videodisc controller. The computer serves as the controller for the workstation and as a music score display. The videodisc is used to provide detailed visual and verbal instructions, the keyboard is the student's input device and the synthesizer allows the student to hear what he or she has played.

In order to evaluate the student's performance, turn music score pages at the appropriate times and accompany a student's playing, the system has to be able to recognize and evaluate keyboard input. The MIDI interface reports the notes played, but the heart of the system is the pattern-matcher. The matcher allows the system to identify where the student is in a score, even though he may be playing at varying rates of speed and playing many wrong notes. The information gathered by the pattern-matcher is used to:

---

[2]*MIDI is an industry standard musical instrument to computer interface

1. evaluate the student's performance;
2. accompany the student, and
3. coordinate the display of the score ("turn pages" on the computer screen).

Although the practicing and teaching systems are separate components, each has similar software building blocks:

1. pattern-matcher - matches music performance with music score in real time;
2. digital recorder - records keyboard performances in digital form;
3. playback - plays recorded keyboard performances;
4. performance evaluator - conducts evaluation of student performance;
5. score display system - handles all score displays;
6. intelligent decision maker - determines the appropriate lessons to execute relative to immediate student input from the keyboard and student performance history;
7. videodisc controller - handles videodisc interactions;
8. accompanist - provides accompaniment for any prerecorded score.

## 4. The Conceptual Architecture

The system provides direct instruction in its teaching-mode and passive criticism in the practicing-mode. In the teaching mode, the system actively teaches new concepts, evaluates student's performance and directs the student's progress. The practicing system provides passive visual criticism to guide the student in evaluating and correcting his or her errors without without interruption during play.

### 4.1. Piano Tutor as an Intelligent Piano Teacher

In its teaching mode, Piano Tutor instructs beginning piano students in basic performance skills. At the outset we intend to develop the curriculum for approximately one year's worth of piano instruction. The schematic below shows a high-level view of the teaching system.

The system is composed of four parts: working memory, lesson knowledge base, lesson planner expert and lesson inference engine. Dynamic memory, static memory, overall lesson execution and individual lesson execution are separated into these respective parts. The lesson planner is in charge of overall instructional decision-making. The lessons and lesson inference engine make low-level instructional decisions for minor corrections. This component architecture is intended to make the system easily modifiable, allowing us to study and refine the instructional environment systematically. The diagram below shows the system components and how they correspond to each other.

#### 4.1.1. Lesson Teacher Inference Engine

The lesson teacher inference engine takes one lesson and executes it in a systematic way. The inference engine has no predefined notion of what the knowledge base contains. How the lessons are run will be discussed in depth later on.

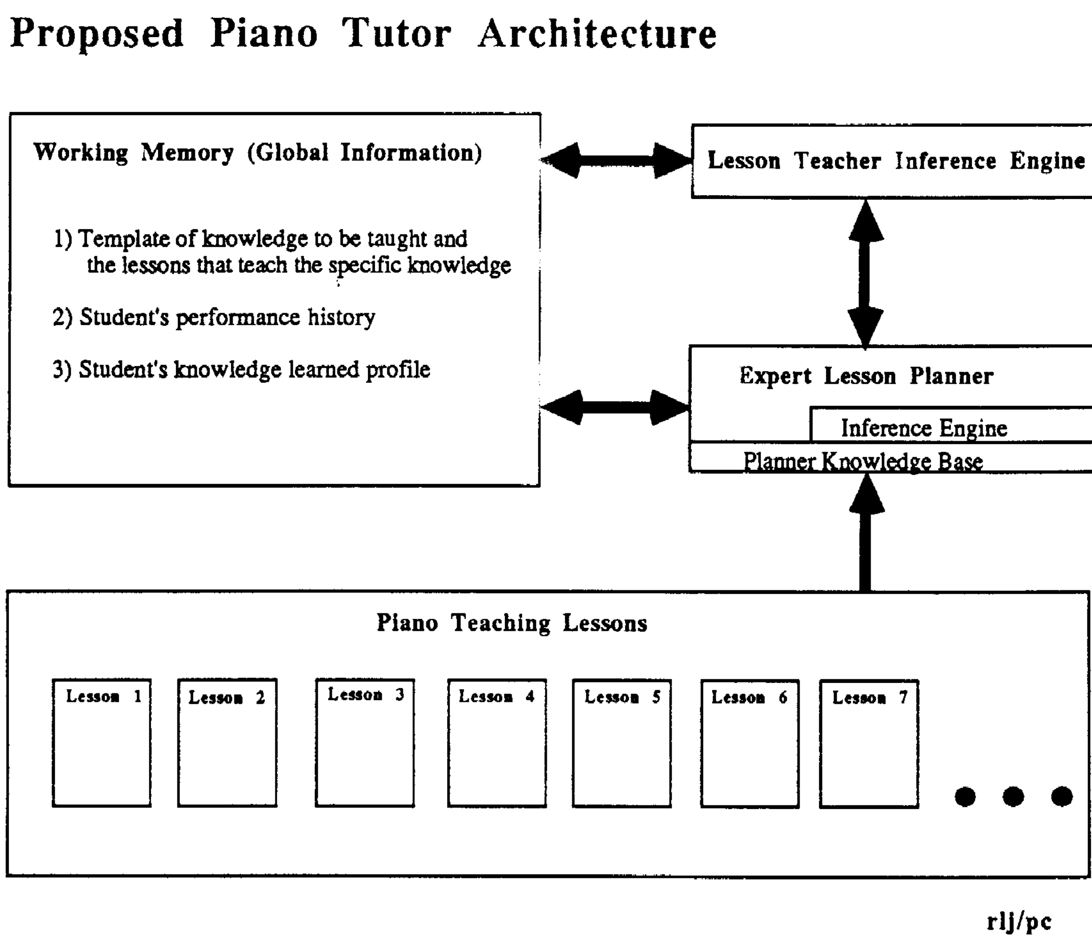


**Figure 4-1:** The Teaching System Architecture

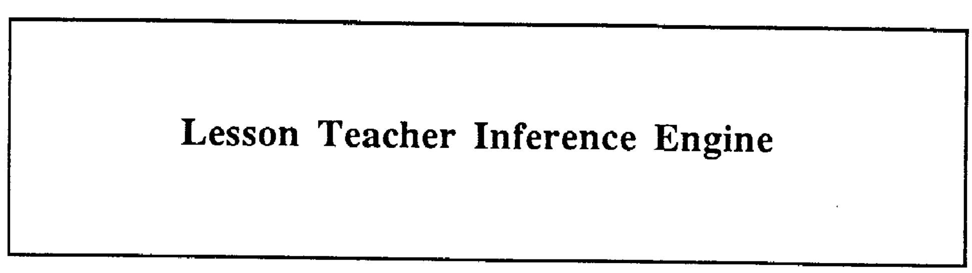


**Figure 4-2:** Lesson Teacher Inference Engine

### 4.1.2. Lesson Knowledge Base

The lesson knowledge base contains individual lessons programmed for specific content and presentation. Each lesson is a separate entity and can address remediation locally when small problems arise. If a serious problem is discovered, the system may conclude that the student was mistakenly given credit for having fulfilled prerequisites for the lesson and that missing skills must be taught before going

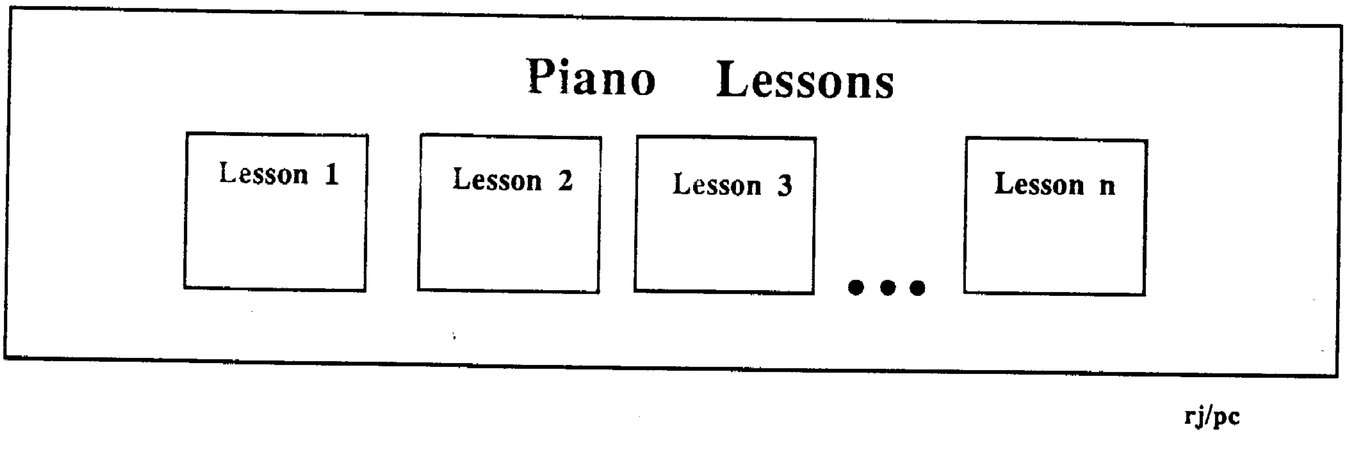


Figure 4-3: Lesson Knowledge Base

any further. The lesson is then is then terminated and the lesson planner determines the appropriate lesson for remediation. Any terminated lesson may be activated when the student's performance improves. A description of the student's performance at the end of each lesson adds information to the student's profile and history.

#### 4.1.3. Expert Lesson Planner

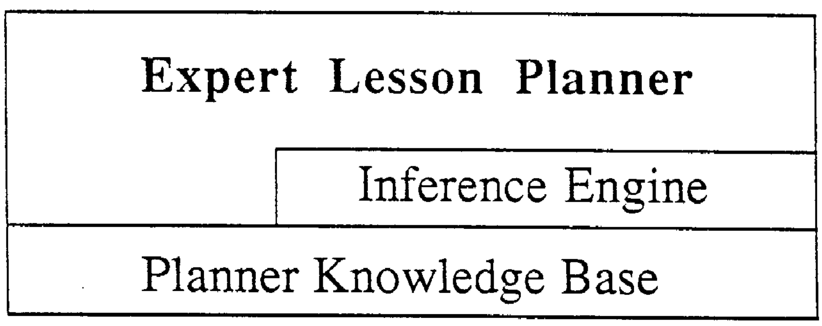


Figure 4-4: Expert Lesson Planner

The expert system lesson planner schedules lessons to be taught. It determines which lesson is to be taught by referring to its expert knowledge base evaluating: the student's profile, the performance evaluation that the current lesson returned, the default order of the lesson presentations and the location in the system of skills and knowledge to be taught. The lesson planner knowledge base knows nothing about underlying lessons, it simply manages which lessons fire according to incoming evaluation reports from active lessons. The system also updates the student's profile and history.

#### 4.1.4. Working Memory

Working memory contains dynamically changing information about the student and about the system. This information is the student profile. It includes the history of the student's performance and information about what the system can teach the student and where appropriate lessons can be found. As the student learns more, these internal models are updated.

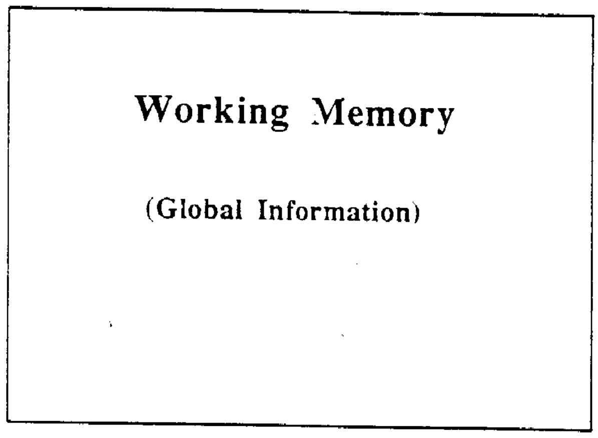


Figure 4-5: Working Memory

### 4.1.5. Presentation Frames

Lessons in the Lesson Knowledge Base must present a large amount of information to the student using various modes including speech, music, graphics, text, and video. To make multimedia presentations more manageable, each presentation of lesson material is encoded into a structure called a presentation frame.

Rather than presenting material directly, lessons call upon presentation frames to do the work. This makes it easy to fine-tune lesson materials; instead of reprogramming a lesson, one can simply edit the corresponding presentation frame. Presentation frames also allow some flexibility in the mode of presentation. For example, the frame can contain text and graphics to be presented in case a videodisc player is not attached to the computer system.

A typical presentation frame might indicate the following information:

> "On the computer screen, display the score for exercise 23. Then simultaneously show frame 2079 of the videodisc and play prerecorded speech from file "exercise23intro.speech". Then play videodisc frames 2080 through 2400 with the audio channel. Then print "Ready" above the first note of the score."

At the completion of the presentation, control returns to the lesson which, in this case, will evaluate a student performance. The lesson may call on several other presentation frames to give additional information, encouragement, and evaluations to the student.

## 4.2. The Practicing System

During practice, students usually receive no feedback, except for what can be remembered from the last lesson. The student may not be aware that he is playing wrong notes or rhythm, or may have forgotten specific comments and suggestions. This slows progress and can be a persistant source of frustration to both student and instructor. To assist in making practice time more useful, the Piano Tutor has a piano practicing system.

The practicing system is designed as a passive guide to individual study without intervention. It provides examples of correct performances, evaluates student performance and allows the student to hear his or her own performance through playback. Because this system is intended as an aid to practice, the student has complete control over his or her interaction with the system and has responsibility for interpreting the system's evaluations.

The practicing system can:

1. display scores and play any selected song;
2. evaluate student performance;
3. accompany the student;
4. record and play back the student's performance and
5. play a pre-recorded expert's performance.

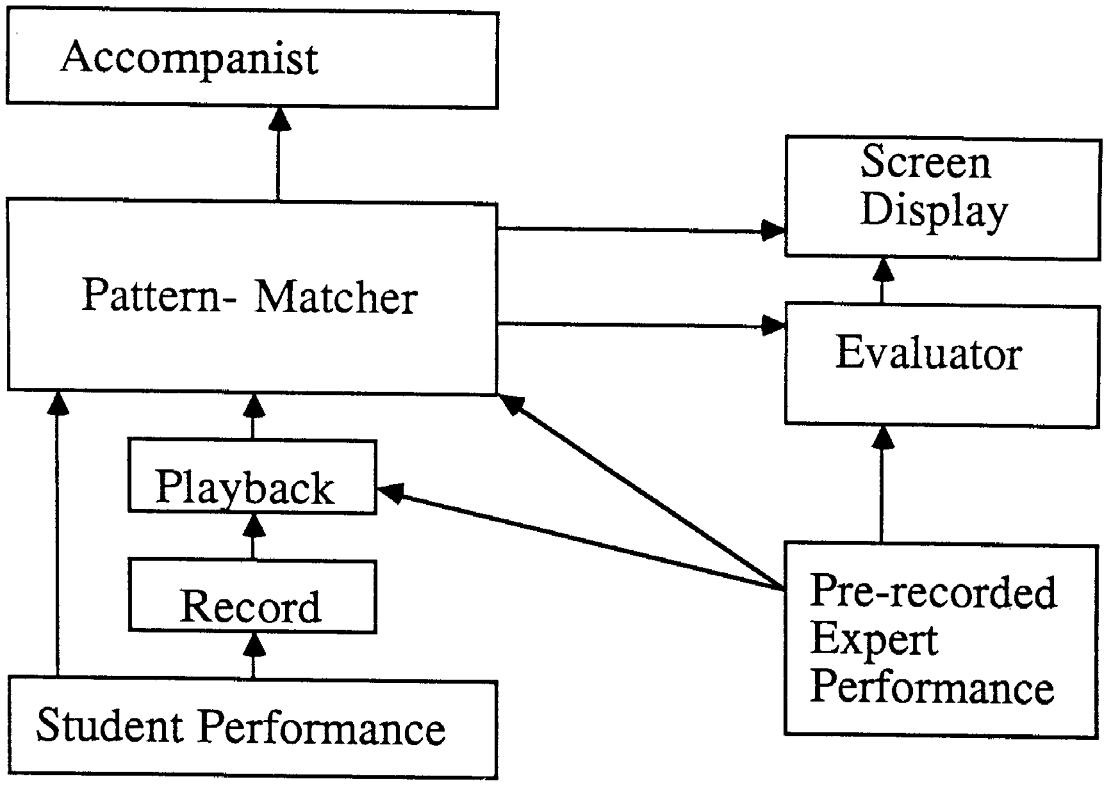


**Figure 4-6:** The Practicer System Component Architecture

The system is activated with the student's performance. Since the student has the option of recording his performance, the record and playback boxes are shown as an alternative path leading out of the student performance box. Whether or not the student chooses to record and playback a performance, keyboard input winds up at the pattern matcher, shown at the center of the diagram. The matcher is the heart of the system, providing processed performance data to the accompanist, screen display and intelligent evaluation modules. The pre-recorded expert performance module provides the scores for accompaniment, screen display and most importantly, the basic information for reference by the intelligent evaluation module.

These tools create an environment for the student to explore and experiment with his or her own playing techniques. Unlike a recording player-piano, the student has ready access to expert evaluation.

## 5. Current Status

The Piano Tutor project has advanced to a stage where we now have a clear picture of the software organization, foundation software has been written, and some preliminary lessons that exercise and test this foundation have also been written. This development is manifest as a demonstration in which a student is led through a simple exercise, complete with remediation if he or she has difficulty. We will now describe in greater detail what is currently implemented.

### 5.1. The Frame System

The Piano Tutor software uses structures called "frames" to store all information and knowledge. Frames are especially convenient for developing artificial intelligence programs. We are using the "Frame Kit" implementation which was developed at CMU and which has been ported to several varieties of Lisp.

### 5.2. The Lesson Interpreter

The center of control in the Piano Tutor is the Lesson Interpreter, whose job it is to select a lesson and deliver lesson material to the student. Since the specific actions of the Lesson Interpreter are controlled by data stored in frames, the Lesson Interpreter is not modified when new lessons are added. The Lesson Interpreter is complete and operational but modifications described in the next section are likely to be needed as our research progresses.

### 5.3. Music Notation

The Piano Tutor can display music notation which includes standard music symbols as well as the ability to highlight notes and turn pages. The notation software in use now is operational but has some limitations. A new system which meets our requirements is available but has not yet been integrated with our other software.

### 5.4. The Performance Interface

Software to record and play musical performance is complete. Only a preliminary evaluation module in use at present, however, and more work is needed here.

### 5.5. Accompaniment and Real Time Evaluation

At present, none of the accompaniment software has been integrated with Piano Tutor. We have designed a method for interrupting student performances when particular problems arise, but this has not been implemented.

### 5.6. Lesson Design

One of the largest components of Piano Tutor is the collection of lessons, encoded as frames, that embody all of the details of teaching piano. We have only a few preliminary lessons at this stage.

### 5.7. Videodisc

Our present system does not use videodisc.

### 5.8. Status Summary

In summary, we have a solid foundation of useable and portable software as well as a considerable amount of experience. The primary things that are lacking in our current system are:

- The ability to deal with multiple simultaneous notes (polyphony and chords), and to accompany student performances.
- Design and implementation of an integrated set of lessons.
- The incorporation of videodisc for presentation of lesson materials.
- The use of a faster workstation and a commercial Lisp implementation.

## 6. Research Opportunities

### 6.1. ICAI and Psychomotor Skills

The Piano Tutor is an instructional program for a complex psychomotor skill. While several existing intelligent teaching systems address cognitive skills, few if any systems have been designed around the substantially different requirements of teaching a highly articulated physical skill. Since the design of curricula for psychomotor skills is a substantially different process from cognitive instructional design, we stand to learn a great deal about the design and development of educational and training systems for an activity that involves intricate hand/eye coordination.

### 6.2. Videodiscs

Videodiscs and compact discs (CD's) represent a revolution in mass-storage technology and rapid access to information. When used to store analog video in constant angular velocity (CAV) format, a videodisc can hold up to 30 minutes of high-quality video with stereo audio on each side. Under computer control, a videodisc player can access any image or motion sequence on a side in less than 1.5 seconds (worst case), with typical access well under a second. This combination of high-quality video and rapid access has been used to advantage in many CAI (Computer-Aided Instruction) systems, particularly in business and government, causing videodisc equipment and production costs to drop, in turn making them more attractive for educational applications and research.

Through a combination of live-action sequences on the videodisc and several hours of the teacher's voice, we explore how effective we can be in providing meaningful and effective feedback through these media. We believe our development strategy solves a number of problems and anticipates future interactive video systems. In particular, we use an expert system to control the videodisc. This avoids tedious programming of limited videodisc controllers, and the use of a powerful computer will allow us to edit voice and music materials interactively.

## 7. Work Plan

We expect to spend roughly half of a year extending Piano Tutor to handle polyphonic scores and moving Piano Tutor to a Macintosh II workstation. Lesson design can proceed in parallel with this effort. We will then spend half of a year programming these lessons and extending Piano Tutor as necessary. As time allows, we also plan to experiment with the addition of a videodisc to the system. These tasks, roughly in order of completion, are described in greater detail below. The time allotted to each task is given in parenthesis.

### 7.1. Porting to the Macintosh II

At present, the best machine for further development is the Macintosh II, which allows us to run real-time software, has a reasonable implementation of Common Lisp, and is fairly fast. Because the Macintosh II is similar to the Macintosh Plus (our current machine) porting will be limited to:

1. Install common Lisp and port Frame Kit to the Macintosh (one week).
2. Build interfaces to our existing library of C programs so that they will be available from within Common Lisp (four weeks).
3. Build additional user interface routines for text and graphics displays and menu selection as necessary (two weeks).

### 7.2. Polyphony

1. Design techniques for the representation and analysis of errors in polyphonic performances (four weeks).
2. Implement polyphonic error representation and analysis including real-time interruption (four weeks).
3. Modify the lesson interpreter's remediation code to handle knowledge about polyphonic errors (two weeks).
4. Modify accompaniment code to handle polyphonic scores and integrate with Piano Tutor (six weeks).

### 7.3. Videodisc

Build a software interface to a commercial videodisc player. Extend the presentation frame interpreter to recognize and handle videodisc presentations (three weeks).

### 7.4. Lesson Design

Lesson design will proceed in parallel with all of the tasks listed above (six months).

### 7.5. Lesson Implementation

Once the primary components of the Piano Tutor are in place, the implementation task becomes a matter of programming small units of instruction (lessons). A typical lesson implementation includes:

1. Create presentation material in the form of speech, text and music notation. Enter descriptions of this data as presentation frames (one day).
2. Program an analysis of student behavior which invokes remediation lessons as necessary and updates the internal model of the student's knowledge. This task is largely a matter of providing a set of conditions to test for and filling in appropriate actions for each condition

(two days).

3. Evaluate and debug lessons through actual trial use. Tuning the lessons according to the knowledge of our experts is likely to take many iterations and experimentation (three days).

Thus, the average time to implement one lesson is roughly six days and we can expect to implement about 20 lessons in six months. We expect these lessons to cover two to four months of piano instruction.

### 7.6. Incorporating the Videodisc

While our programmer is concentrating on programming lessons, we plan to design and shoot some video presentation material. While video is probably the most time-consuming material to prepare, incorporating video into the Piano Tutor and integrating it with lessons should be very simple once we have the material.

### 7.7. Product Transition

The next 20 lessons should be easier than the first, but this is offset somewhat by the fact that the student knows more on his 21st lesson than on his first. A rough guess is that we can continue producing 20 lessons per year, and perhaps 60 lessons, or three years of effort, will be required to implement a year of piano instruction.

At that point, adding new lessons should be a routine matter, and there will not be a significant amount of further work appropriate for a university aside from further curriculum design and refinement. Electronic Arts, the largest publisher of music software in the U.S. has expressed interest in our work and we also have close ties to NeXT, Inc., Yamaha, and Commodore, all of which make and sell computers and have a keen interest in music education. We hope that with a relatively complete prototype, one or more of these firms will be interested in further development and marketing of Piano Tutor systems. This will take an additional two to three years. By then, new technologies will probably replace videodisc and be low in cost if not commonplace in homes for entertainment. A total system cost, including a synthesizer, of a few thousand dollars is not unreasonable to expect.

## 8. Deliverables

At the end of one year, we should have a working tutor system that can deliver around 20 lessons of piano instruction. The system will present instruction using music notation, text, voice, synthesized music and perhaps video. We expect to have used the system to teach real students, but only with close supervision so that we can intervene when bugs arise or when the Piano Tutor's knowledge is incomplete. At the end of one year, we expect there will still be gaps in the system's knowledge, but we believe it will be useful to "sympathetic" users.

## 9. Summary

The Piano Tutor can provide us with many interesting avenues of research in its design and development. By integrating newly emerging technologies: real-time music recognition technology, intelligent teaching systems and high-capacity optical storage media, we stand to learn a great deal about each of these and to be able to share that knowledge with others. Our basic objective is to create a useful tool for teaching basic piano skills.

With the technologies and expertise availaole to us, we have an opportunity to capture some of the basic essence of the master/apprentice relationship in formal terms. The exercise of codifying the rules for this kind of instruction permits us to study an important and effective kind of educational exchange with rigor and precision. Understanding the basic constituents of instruction would be an accomplishment in and of itself - a goal which has eluded educational scientists to this day. At the very least, this project should contribute a number of valuable practical insights.