J. EDUCATIONAL COMPUTING RESEARCH, Vol. 42(2) 219-239, 2010

A COGNITIVE TUTOR FOR GENETICS PROBLEM SOLVING: LEARNING GAINS AND STUDENT MODELING*

ALBERT CORBETT, LINDA KAUFFMAN, BEN MACLAREN, ANGELA WAGNER, ELIZABETH JONES

Carnegie Mellon University

ABSTRACT

Genetics is a unifying theme of biology that poses a major challenge for students across a wide range of post-secondary institutions, because it entails complex problem solving. This article reports a new intelligent learning environment called the Genetics Cognitive Tutor, which supports genetics problem solving. The tutor presents complex, multi-step problems and is constructed around a *cognitive model* of the knowledge needed to solve the problems. This embedded cognitive model enables the tutor to provide step-by-step assistance, and to maintain a model of the student's problem-solving knowledge. The tutor consists of 16 modules with about 125 total problems, spanning five general topics: Mendelian inheritance, pedigree analysis, genetic mapping, gene regulation, and population genetics. This article reports two evaluations of the tutor. A pretest/posttest evaluation of student learning gains for individual tutor modules across multiple colleges and universities yielded average gains equivalent to almost two letter grades, and the accuracy of student modeling in predicting students' test performance was empirically validated.

*This research was funded by U.S. Department of Education Fund for the Improvement of Post-Secondary Education grants P116B020344 and P116B060106, by a Howard Hughes Medical Institute Professorship to Elizabeth W. Jones, and by National Science Foundation grant DUE-0231219.

doi: 10.2190/EC.42.2.e
http://baywood.com

This article describes a new intelligent learning environment called the Genetics Cognitive Tutor that supports student problem solving in genetics. This Cognitive Tutor poses rich, multi-step problem-solving tasks to students and provides the individualized advice students need to succeed. The tutor is constructed around a *cognitive model* of the knowledge required to solve the problems. This cognitive model is a type of embedded expert system that can solve the problems posed to students in the many ways that students solve them. The cognitive model is used to follow each student's solution path through complex problems, enabling the tutor to provide step-by-step accuracy feedback, to provide advice as needed, and to maintain a model of the student's problem-solving knowledge based on student performance. In the following sections, we provide an overview of the tutor, describe the cognitive model that underlies its performance, and report empirical evaluations of student learning gains and of the predictive validity of the tutor's student model.

Genetics is a fundamental unifying theme of biology and a key component of scientific literacy (AAAS, 1993; Garton, 1992; Lewis & Wood-Robinson, 2000; NRC, 1996). Advances in genetics underlie key areas in 21st century technology, science, and industry, from forensic DNA analyses to detecting and understanding the causes of cancer. These advances require a workforce capable of contributing to these fields, and also raise complex ethical and legal issues that will require a comprehensive understanding of genetics. However, research suggests that students are currently not succeeding at developing the deep understanding of genetics necessary to participate in these activities (Garton, 1992; Lewis & Wood-Robinson, 2000).

Problem solving is heavily emphasized in genetics (Smith, 1988). Genetics problem solving is characterized by abductive reasoning, a complex type of reasoning used in real-world analysis of genetics data (cf. Papatheodorou, Kakas, & Sergot, 2005; Zupan, Bratko, Demšar, Juvan, Curk, Borštnik, et al., 2003). In contrast with deductive hypothesis testing, abductive reasoning in genetics starts with a set of observations and reasons backwards to infer properties of the genetics processes that produced the data (e.g., whether a trait expressed in a family pedigree chart is dominant or recessive). In addition to the challenge of abductive reasoning, genetics problem solving relies on successful use of mathematical skills, particularly reasoning about probability (Cavallo, 1996; Smith, 1988), but including algebraic modeling and symbol manipulation, and proportional reasoning. The dual learning goals of genetics problem-solving activities are to develop the big ideas that underlie genetics and to develop students' reasoning skills in various problem situations.

THE GENETICS COGNITIVE TUTOR

In response to this challenge, we have developed a Genetics Cognitive Tutor to support students in completing and learning from complex problem-solving tasks. Cognitive Tutors pose complex problems and employ a cognitive model of relevant domain knowledge and problem-solving skills to provide the individualized help students need to complete the problems (Anderson, Corbett, Koedinger, & Pelletier, 1995). Successful Cognitive Tutors have been developed for programming and mathematics that yield large improvements in student learning (Anderson et al., 1995; Koedinger, Anderson, Hadley, & Mark, 1997; Koedinger & Corbett, 2006). Cognitive Tutors have been shown to speed learning by as much as a factor of three (Corbett & Anderson, 2001) and to yield an achievement effect size of about one-standard deviation compared to conventional instruction (Anderson et al., 1995; Corbett, 2001). This is about twice the achievement effect size of typical human tutors (Cohen, Kulik, & Kulik, 1982) and conventional computer-based instruction (Kulik, 1994), and about half the effect of the best human tutors (Bloom, 1984; Kulik, 1994).

The Genetics Cognitive Tutor consists of 16 modules, spanning five general topics: Mendelian inheritance, pedigree analysis, recombination and genetic mapping, gene regulation, and population genetics. The 16 individual modules are displayed in Table 1. There are a total of about 125 problems in the curriculum, with an average of about 25 problem-solving steps per problem. The use of a genetics problem-solving cognitive model to interpret student actions and provide advice, as well as to model student knowledge and individualize the curriculum, distinguishes the Genetics Cognitive Tutor from other educational software packages that have been developed for genetics. Some other genetics software packages are helpful, rich simulation environments that allow students to breed organisms across generations, conduct experiments, and reason about inheritance patterns, but provide no in-context instruction nor assessment. Other software environments provide accuracy feedback and in-context instruction, but only for multiple-choice or short-answer questions. These other packages do not combine rich problem-solving activities with feedback on individual steps and context-sensitive instruction.

An Example Genetics Cognitive Tutor Problem-Solving Task

Human pedigree analysis is a typical example of genetics problem solving. In pedigree analysis problems students are asked to examine the pattern of individuals in a family tree who are affected by a disease trait or unaffected by the trait, and to reason about whether that disease trait is dominant or recessive, and whether it is transmitted on the X-chromosome or on one of the 22 autosomal chromosomes. In the case of recessive traits, the student may also be asked to reason about the probability that unaffected individuals are carriers of the disease. Figure 1 displays the tutor screen near the beginning of a problem selected from the Pedigree Analysis and Carrier Probabilities module. These problems display the pedigree chart of a family in which some individuals display a rare trait. Circles represent females, squares represent males, and the darkened individuals (II-2

Table 1. Sixteen Genetics Cognitive Tutor Modules Spanning Five Major Topics

Major topics	Modules
Mendelian Transmission	1. Mendelian Transmission: Predictive Modeling
(Modules 1-3)	2. Mendelian Transmission: Experiment Design and Interpretation
	3. Gene Interaction and Epistasis
Pedigree Analysis	4. Basic Pedigree Analysis
(Modules 4-7)	5. Basic Pedigree Analysis with Student Explanations
	6. Pedigree Analysis & Carrier Probabilities
	7. Pedigree Analysis & Carrier Probabilities with Explanations
Recombination and Genetic	8. Recombination and Segregation: Forward Modeling
Mapping (Modules 8-11)	9. Three-Factor Cross
	10. Tetrad Analysis
	11. Hfr Time of Entry in *E. coli*
Gene Regulation	12. Gene Regulation 1: Predictive Modeling
(Modules 12-13)	13. Gene Regulation 2: Experiment Design and Interpretation
Population Genetics	14. Hardy Weinberg Equilibrium
(Modules 14-16)	15. Identifying Hardy-Weinberg Equilibrium
	16. Departures from Hardy-Weinberg Equilibrium

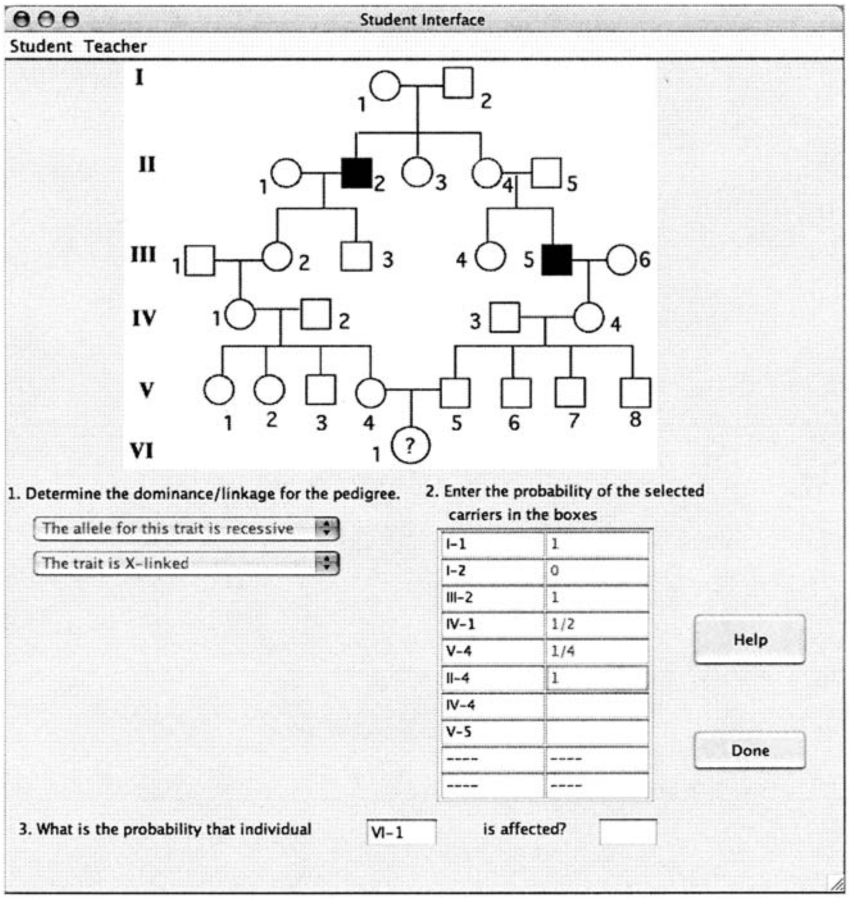

Figure 1. The Genetics Cognitive Tutor Pedigree Analysis/ Carrier Probability problem-solving interface partway through a problem.

and III-5 in this diagram) are affected by the disease. Finally, the circle with a question mark at the bottom, individual VI-1, represents the potential daughter of two individuals in the fifth generation of the pedigree; the question mark indicates that her phenotype (whether or not she has the disease) is unknown. The student has three problem solving goals:

1. The first steps are to determine whether the rare trait is dominant or recessive and whether the gene is X-linked or is autosomal, from the pattern of affected and unaffected individuals.

2. Next, students calculate the probability that specified unaffected individuals (listed in the table on the right side of the screen) are carriers of the trait.
3. Finally, students determine the probability that the unobserved hypothetical offspring VI-1, will be affected by the trait.

In Figure 1, the student has indicated that the trait is recessive and X-linked on the lower left side of the screen, and filled in the probabilities that six of the individuals are carriers in the table at the lower right. The tutor provides immediate accuracy feedback on each step and the student can ask for problem-solving advice at any step, so students always reach a complete and correct problem solution. If the student asks for advice in the problem-solving state displayed in Figure 1, the tutor focuses on the probability that individual IV-4 is a carrier and provides the first message in Figure 2. This message suggests a general abductive reasoning strategy to apply in the current context—how to begin reasoning about an unaffected female in an X-linked pedigree. If the student requests more advice, the second message focuses in on the relevant empirical evidence in this context. A third help request elicits instruction on the relevant underlying genetic transmission process and, finally, a fourth request elicits the final conclusion on the probability that the female IV-4 is a carrier.

THE COGNITIVE MODEL

As in previous Cognitive Tutors, each Genetics Cognitive Tutor module employs a *cognitive model* of genetics problem-solving knowledge to provide the step-by-step assistance described above. The cognitive model is a type of expert system embedded in the tutor module that can solve the problems posed to students

(1) To determine the probability that an unaffected female is a carrier of an X-linked recessive trait, the first thing to consider is the genotype of her parents.

(2) Individual IV-4 has an affected father. What can you conclude about the probability that IV-4 is a carrier?

(3) Her father must have the disease allele on his single X chromosome and he must transmit that X chromosome to his daughter.

(4) Since IV-4 must inherit the disease allele from her father, the probability that she is a carrier is 1.

Figure 2. An example of problem-solving advice with four levels of help for one step in the problem displayed in Figure 1.

in the many ways that students solve them. This cognitive model is grounded in cognitive psychology: each tutor module employs a model derived from a cognitive task analysis of the domain knowledge and reasoning strategies students are learning and applying in problem solving. The cognitive model is employed for two purposes in a Cognitive Tutor, to model and to adapt to the individual student. First, as described above, the model is employed to interpret the students' problem-solving actions and to provide feedback and advice, in a process called *Model Tracing*. Second, it is employed to monitor each student's growing problem-solving knowledge, in a process called *Knowledge Tracing*.

The Cognitive Model for Pedigree Carrier Probabilities

The cognitive model is based on ACT-R theory (Anderson, 1993; Anderson & Lebiere, 1998), a unified theory of the nature, acquisition, and use of human knowledge. ACT-R assumes that problem-solving knowledge is represented in the form of a set of goal-specific if-then knowledge chunks called *cognitive rules*. There are two general sets of rules in the cognitive model for the Pedigree Analysis and Carrier Probabilities unit. One set reasons about the dominance and linkage of traits. An "English translation" of one of these rules is:

- If an affected individual (e.g., II-2) has unaffected parents, Then the trait must be recessive.
  (Since the individual must have inherited at least one allele for the trait from at least one parent, and a parent with even a single dominant allele would be affected).

A second set of cognitive rules reasons about the probability that unaffected individuals in the family carry an allele for the trait and the probability that the hypothetical descendant will be affected. Some example rules for an X-linked recessive trait (the type displayed in Figure 1) include:

- If a male is unaffected (e.g., I-2), Then he cannot be a carrier.
  (Since males have a single X chromosome, a male will be affected by a single X-linked disease allele and cannot be a carrier).
- If an unaffected female (e.g, III-2) has an affected father, Then she must be a carrier.
  (A father always passes his single X-chromosome to his daughters, along with the X-linked allele of the rare trait).

Model Tracing

In Model Tracing, the cognitive model runs in step-by-step synchrony with the student. At each problem-solving step, the student's problem-solving action is compared to all the problem-solving actions the model is capable of generating for the step. As with effective human tutors, cognitive tutor feedback is brief and focused on the student's problem-solving context. If the student action matches the action of an applicable cognitive rule, it is simply accepted by the tutor. If the student action matches no applicable rule, it is rejected and flagged as incorrect. The cognitive model also includes some *buggy rules*, which represent common student misconceptions and if an action matches a buggy rule, it is flagged as incorrect and the tutor displays a brief just-in-time error message in the hint window. The tutor does not automatically provide advice on an incorrect problem-solving step; instead students have the opportunity to reflect on and correct their own mistakes. However, if the student asks for advice, the cognitive model is also employed to provide problem-solving advice: An applicable rule is identified and associated help message templates are employed to construct context-specific advice. As indicated earlier, there are generally three or more levels of advice available for each problem-solving step. These messages describe an appropriate goal and provide advice on satisfying the goal, with successive messages providing more and more specific advice on solving the goal in the current context.

Knowledge Tracing

In Knowledge Tracing the tutor monitors the student's growing knowledge during problem solving, maintaining an estimate of the probability that the student has learned each of the cognitive rules in the cognitive model, based on the student's problem-solving performance. Each problem-solving step represents an opportunity to apply some rule in the underlying cognitive model. Note that some rules may apply multiple times within a single problem. For example, each of these two cognitive rules applies twice in the problem displayed in Figure 1:

- If an unaffected female has an affected father, then she must be a carrier.
- If an unaffected female has an unaffected mother and an unaffected father and no affected male descendents, then the probability of being a carrier is 1/2 the probability that her mother is a carrier.

Note also that two applications of the same cognitive rule generally do not yield the same outcome. For, instance, the second rule immediately above applies to individuals IV-1 and V-4 in Figure 1, but the probability value generated by the rule is different for the two individuals.

At the conclusion of each problem-solving step, the tutor updates its estimate of the probability that the student has learned an applicable rule. This updated probability estimate is the sum of two probabilities:

1. A revised estimate of the probability that the rule was already in the learned state, given the new evidence (a correct or incorrect response or help request); and

2. the probability the student learned the rule at this opportunity if the student did not already know the rule. See Corbett and Anderson (1995) for computational details. Knowledge Tracing in turn is employed in cognitive tutors to implement *Cognitive Mastery*, a process in which the tutor continues selecting appropriate problems based on the student's knowledge state, until the student has learned each rule to a criterion level (Corbett & Anderson, 1995).

EMPIRICAL EVALUATION: LEARNING GAINS

Genetics Cognitive Tutor modules have been piloted in 12 colleges and universities around the country. These include a broad range of post-secondary institutions, including public and private institutions, liberal arts institutions and research universities, and minority-serving universities. The tutor has been piloted in seven different courses across these institutions, ranging from introductory biology to upper level genetics courses. For each course, the instructor selected between two and six tutor modules that both matched the course content and fit course time constraints. Students in the courses completed the Cognitive Tutor problems either as in-class activities in section or lab meetings, or as homework assignments.

Across all these contexts, we have conducted and analyzed 36 pretest-posttest evaluations of individual modules. Each of these single-unit evaluations consisted of:

1. completion of a proctored paper-and-pencil pretest (approximately 20 minutes);
2. completion of the Genetics Tutor problem solving module (approximately 1 hour); and
3. a proctored paper-and-pencil posttest (approximately 20 minutes).

Students worked through a fixed problem set in these modules. The paper-and-pencil tests consisted of problems similar to the Cognitive Tutor problems and two paper test forms were constructed for each module. The tests were designed to be very challenging for students, in order to avoid ceiling effects in measuring learning gains. Each form serves as the pretest for half the students, who then switch to the other form for the posttest, so that the pretests and posttests are matched across the full set of students, but for each student the pretest and posttest are different. As displayed in Table 2, the tutor module was used as an in-class activity in 22 of the evaluations and as a homework assignment in the other 14 evaluations. A total of 1050 students participated in these evaluations.

Average student pretest and posttest performance (percent correct) in each of these 36 evaluations is displayed in Table 2, grouped by main genetics problem-solving topics. Overall, the Cognitive Tutor modules proved very effective. Across the 36 evaluations, students averaged 43% correct on the pretest and 61%

Table 2. A Summary of 36 Single-Module Evaluations of Student Learning Gains

Genetics topic	Students	Context	Pretest (% Correct)	Posttest (% Correct)	Gain (Percentage points)	*t*
I. Transmission						
Gene Interaction	12	homework	20	61	41	9.55***
Gene Interaction	11	homework	22	46	24	3.24***
Gene Interaction	37	homework	31	59	28	7.07***
Gene Interaction	10	homework	48	80	32	3.13**
Gene Interaction	56	homework	30	60	30	8.04***
Mendelian Transmission 2	23	homework	57	55	–2	–0.39
II. Pedigrees						
Basic Pedigree Analysis	15	in-class	29	40	11	2.07*
Basic Pedigree Analysis	15	in-class	37	46	9	1.42
Basic Pedigree Analysis	24	in-class	39	47	8	1.93*
Basic Pedigree Analysis	37	in-class	46	67	21	4.96***
Basic Pedigree Analysis	20	homework	52	59	7	1.50
Basic Pedigree Analysis	10	in-class	56	53	–3	–0.50
Basic Pedigree Analysis	20	in-class	69	80	11	2.90***
Basic Pedigree/Expl	71	in-class	49	59	10	4.18***
Carrier Probabilities	44	homework	28	63	35	9.04***
Carrier Probabilities	21	homework	40	53	13	2.80**
Carrier Probabilities	43	in-class	63	81	18	5.63***
Carrier Probabilities	48	in-class	58	75	17	7.33***

III. Recombination/Mapping						
Recomb & Segregation	18	in-class	17	29	12	1.96*
Recomb & Segregation	23	homework	24	30	6	0.92
Recomb & Segregation	19	in-class	35	63	28	4.00***
Recomb & Segregation	11	in-class	36	67	31	3.85***
Three Factor Cross	45	in-class	17	69	52	13.09***
Three Factor Cross	11	homework	25	22	–3	–0.59
Three Factor Cross	88	in-class	34	75	41	14.57***
Three Factor Cross	51	in-class	44	71	27	7.67***
Three Factor Cross	10	homework	45	61	16	3.31***
Three Factor Cross	63	homework	48	59	11	5.01***
Three Factor Cross	16	in-class	52	59	7	2.32**
Time of Entry	38	in-class	47	69	22	8.68***
Time of Entry	35	in-class	50	90	40	9.38***
Tetrad Analysis	11	in-class	69	86	17	4.28***
IV. Gene Regulation						
Gene Regulation 2	18	in-class	63	68	5	1.23
Gene Regulation 2	18	homework	67	66	–1	–0.12
Gene Regulation 2	40	in-class	68	77	9	4.95***
V. Population Genetics						
Population Genetics 2	18	in-class	33	62	29	4.78***
OVERALL	1050		43	61	18	

*$p < 0.10$; **$p < 0.05$; ***$p < .01$.

correct on the posttest and this average gain of 18 percentage points is the equivalent of almost two letter grades. Thirty-two of the 36 evaluations led to positive learning gains, and this overall result is reliable by a simple sign test. A *t*-test was performed on each of the 36 individual evaluations. As indicated in Table 2, 26 of the evaluations yielded reliable learning gains, 2 evaluations yielded a marginally reliable gain, and 8 of the evaluations did not yield reliable learning gains. These evaluations are conducted after the instructor has finished lecturing on the topic, so these gains are in addition to the learning that resulted from traditional instruction. The existing courses do not already contain comparable problem-solving curricula, so these learning gains directly represent the impact of the Cognitive Tutor related to current course practice. These evaluations serve both to validate the general effectiveness of the modules and to identify individual modules that are candidates for further improvement.

There is a wide range in learning gains across the 36 evaluations. Gains exceeded 30 points for 8 evaluations, while for another 8 evaluations there were no reliable gains. With multiple factors varying in an unsystematic fashion across the evaluations, we cannot draw strong conclusions about the possible factors that affect the magnitude of learning gains, but one relationship emerges from the table. Across the 36 evaluations, the size of the learning gain is inversely correlated with average pretest score, $r = -0.45$, $t(34) = -2.90$, $p < .01$. This correlation suggests that across the variations in student populations and topics, the tutor modules tend to yield larger learning gains for the topics that the students initially find more challenging.

Table 3 shows the average learning gains across the 36 evaluations for each of the 11 individual topics. A reliable inverse correlation of pretest score and learning gain is again observed across these 11 topics, $r = -0.60$, $t(9) = -2.27$, $p < .05$. Since the number of evaluations per topic varies, any conclusions are preliminary, but this correlation suggests that the tutor is more helpful specifically for the topics that are initially most challenging, collapsing across student populations.

While these correlations of pretest scores and gain scores are suggestive, they can arise in part through regression to the mean, resulting from test measurement error. To derive more objective measure of topic difficulty, we counted the total number of problem-solving steps in a typical problem in each topic. These steps can include typed inputs and menu-based entries, and we separately counted just the number of steps with typed inputs, since these inputs are more challenging. The total number of steps and total number of typed inputs for each topic are displayed in the two right columns of Table 3. Across the 11 topics, the correlation of learning gains with total number of problem-solving steps is marginally reliable, $r = 0.53$, $t(9) = 1.87$, $p < .10$, while the correlation of learning gains with total number of typed student inputs is reliable, $r = 0.63$, $t(9) = 2.46$, $p < .05$. These correlations provide converging evidence that the Cognitive Tutor Modules are most effective for the topics students find most challenging.

Table 3. A Summary of Average Student Test Performance and Number of Problem-Solving Steps across 11 Genetics Topics

Genetics topic	Evaluations	Pretest	Posttest	Gain	Steps	Typed steps
Gene Interaction	5	30.2	61.2	31.0	41	26
Time of Entry	2	48.5	79.5	31.0	42	42
Population Genetics 2	1	33.0	62.0	29.0	23	19
Three Factor Cross	7	37.9	59.4	21.6	24	20
Carrier Probabilities	4	47.3	68.0	20.8	11	9
Recomb & Segregation	4	28.0	47.3	19.3	35	35
Tetrad Analysis	1	69.0	86.0	17.0	13	7
Basic Pedigree/Expl	1	49.0	59.0	10.0	7	0
Basic Pedigree Analysis	7	46.9	56.0	9.1	2	0
Gene Regulation 2	3	66.0	70.3	4.3	38	19
Mendelian Transmission 2	1	57.0	55.0	–2.0	9	6

EMPIRICAL EVALUATION: KNOWLEDGE TRACING PREDICTIVE VALIDITY

Knowledge Tracing estimates a hypothetical construct—student learning—based on the student's observable performance in Cognitive Tutor problem solving. The probability that a student has *learned* a rule in the cognitive model is not synonymous with the probability the student will *apply the rule correctly*, because the model recognizes that students may "slip" and fail to apply a rule they actually know, or may guess an answer correctly, even if they have not learned an appropriate rule. However, the same performance assumptions (probabilities of slips and guesses) that are employed in Knowledge Tracing calculations can also be employed to predict student performance based on the tutor's model of the student's knowledge. That is, we can predict the probability a student will apply a rule correctly based on the probability that the student knows the rule and the probabilities of slipping and correct guessing. We can validate the tutor's estimates of each student's knowledge by examining how well these Knowledge Tracing estimates predict the student's subsequent paper-and-pencil test performance.

We have begun implementing Knowledge Tracing in Genetics Cognitive Tutor modules and we conducted a preliminary evaluation of its validity in the Pedigree Analysis and Carrier Probability module to examine two questions.

First, we examined how well the student models that Knowledge Tracing constructs predict individual differences among students in test accuracy. Second, collapsing across students, we examined how well Knowledge Tracing predicts overall student accuracy in applying each of the individual cognitive rules on the paper-and-pencil test.

In this empirical evaluation of Knowledge Tracing predictive validity, we analyzed the data for a set of 27 students who completed the seven problems in the Pedigree Analysis Carrier Probabilities module at one pilot university. The Knowledge Tracing algorithm had not yet been implemented in this module, but the Cognitive Tutor collects a keystroke-by-keystroke log file of each student's problem-solving actions, so we can retroactively examine how accurately the Knowledge Tracing algorithm can predict students' application of the cognitive rules in paper-and-pencil posttest performance.

Basic Knowledge Tracing in cognitive tutors employs a set of two learning and two performance parameters for each rule. The two learning parameters are:

1. $p(L_0)$ the probability that students will have learned the cognitive rule prior to the first opportunity to apply it in problem solving; and
2. p(T), the probability that students will learn the cognitive rule at each opportunity to apply the rule in problem solving.

The two performance parameters are:

1. p(S), the probability students will "slip" and make a mistake in applying a rule even though they have already learned it; and
2. p(G) the probability students will "guess" correctly when they have not yet learned a cognitive rule.

In this analysis we evaluated a variation of Knowledge Tracing that also employs four individual difference weights for each student, one weight for each of the four parameters, to estimate differences among students in the four learning and performance parameters. As the first step in validating the Knowledge Tracing algorithm in this module, we submitted the student performance data in the tutor protocol files to a curve fitting program to generate a best-fitting set of learning and performance parameter estimates for each of the rules in the cognitive model, and to generate a best fitting set of individual weights for each of the students. Then we applied the Knowledge Tracing algorithm with these best-fitting estimates to the step-by-step student performance data in the protocol files to estimate the probability that each student knows each rule at the conclusion of problem solving. Finally, we employed these estimates of the probabilities that a student has learned the cognitive rules and the best-fitting slip and guess estimates from tutor problem solving to predict how well each student would perform on the quiz, by predicting the probability of a correct action at each step on the quiz.

Predicting Carrier Probability Test Performance

To calculate Knowledge Tracing predictions for test performance, the first step is to construct a list of the cognitive rules that fire at each of the steps in the test problems. The probability that a student completes a problem-solving step correctly is a function of the probability that the student has learned an applicable cognitive rule. Note that in principle, the algorithm for predicting a student's step-by-step accuracy in problem solving requires that the student remain on a recognizable solution path. In general, when a student makes an uncorrected error in a test environment, we can no longer model the student's current problem-solving state and cannot trace further problem-solving actions. In programming, for example, when a student inserts an unexpected operator in defining a program, it is difficult to know the student's current thinking and, in fact, there may be no path forward to a correct solution.[1] However, reasoning about carrier probabilities has an important property: even if a student makes a mistake in reasoning about the probability that an individual in a pedigree is a carrier, we can still identify the correct cognitive rules to apply for successive descendents in a pedigree (as long as the incorrect answers are numbers in the range 0-1). For example, in Figure 1 the probability that individual IV-1 is a carrier is one-half the probability that her mother III-2 is a carrier. The actual probability that III-2 is a carrier is 1.0 and the probability that IV-1 is a carrier is 0.5. But if the student makes a mistake and assigns III-2 any probability, *p*, between 0 and 1, we can still evaluate whether the student executes the appropriate rule for IV-1, because that rule will assign the probability *p/2*. As a result, we can exhaustively predict step-by-step accuracy for complete problem solutions, even incorrect solutions. This in turn enables us to make use of all the test data to examine how well the Knowledge Tracing model predicts performance for the 14 individual rules in the cognitive model.

Predicting Student Posttest Accuracy

For each of the 27 students, we calculated the probability that the student will apply each successive cognitive rule correctly in the four test problems, given the tutor's model of the student's knowledge. We then calculated the average of all these probabilities as a prediction of the student's overall test accuracy. The scatterplot in Figure 3 displays the accuracy of the Knowledge Tracing model in predicting student performance. The x-axis represents the model's predictions of test accuracy, the y-axis represents students' empirical performance on the test

[1] In predicting individual differences in student accuracy on tests, it is possible to adapt to this by predicting performance at the level of whole problems, rather than at the level of each step. However, this approach does not apply to investigating predictions about individual cognitive rules on tests.

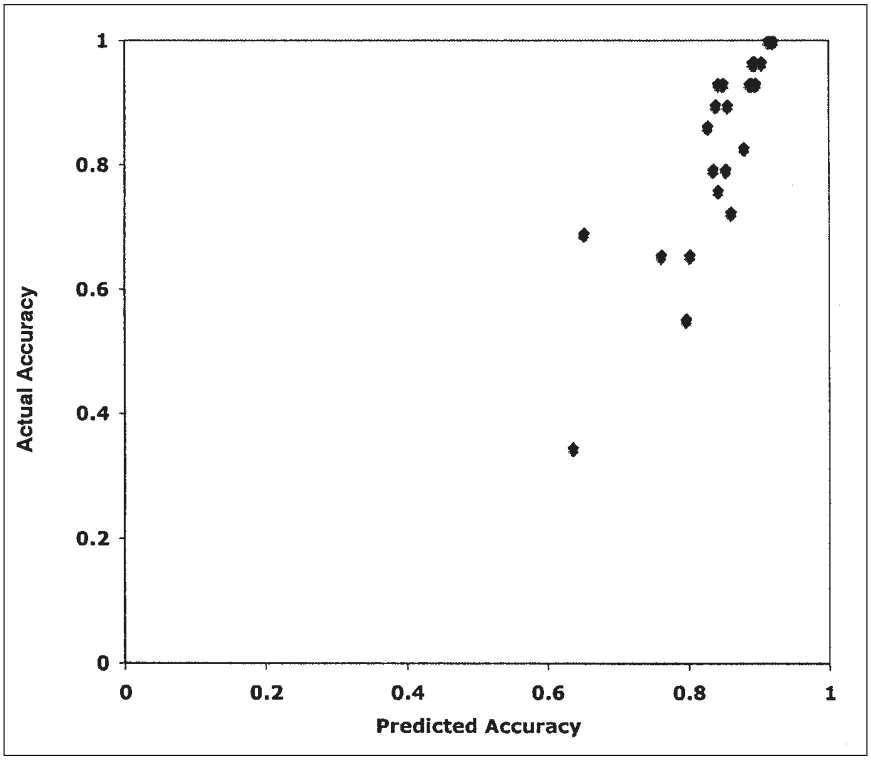

Figure 3. Actual test accuracy plotted as a function of predicted accuracy for 27 students.

and the points represent the 27 students. The model predicts overall group performance quite well; it predicts that students will average 85% correct on the test and the actual average score is 85%. Similarly, the model predicts individual differences among students very accurately. The predicted and actual test scores are highly correlated across the 27 students, $r = 0.85$, $t(25) = 8.07$, $p < .01$.

These results represent an *upper bound* on the predictive validity of Knowledge Tracing, since the analysis employs best fitting parameter values for the set of students in the analysis and employs individual difference weights, but they are a highly promising indication that Knowledge Tracing can be used effectively to implement Cognitive Mastery in the Genetics Cognitive Tutor, that is, to individualize the curriculum, adapting the number of problems to each student's needs. Note that the slope of the regression function is greater than 1. The model slightly underestimates the accuracy of the best students and slightly overestimates the accuracy of the students who are struggling. This suggests that, while less successful students will receive more tutor problems than more successful students and that Cognitive Mastery will close the test performance gap between students, the less successful students may still receive a little less practice than they need in principle, while the more successful students may receive a little more practice than they need.

Predicting Accuracy of Applying Cognitive Rules on the Posttest

In addition to examining how well the Knowledge Tracing model predicts each student's overall test accuracy, we examined how well the model predicts overall student accuracy in applying each of the 14 rules in the cognitive model. In this analysis, we collapsed across the 27 students and for each cognitive rule, we calculated:

1. the average predicted probability of applying the rule correctly across all the opportunities in the test; and
2. the empirical probability that students did apply the rule correctly across all opportunities on the test.

Figure 4 displays a scatterplot of the predicted and actual accuracies for the 14 cognitive rules. Predicted accuracy is on the x-axis, empirical accuracy is on the y-axis, and each of the points represents one of the cognitive rules. Again the model accurately predicts differences among the rules. The predicted and actual scores for the 14 rules are highly correlated, $r = 0.74$, $t(12) = 3.81$, $p < .01$.

Again, the results in Figure 4 represent an upper bound on the predictive validity for Knowledge Tracing, but the model's accurate prediction of student performance is of substantial practical importance in validating the foundations of tutor's Cognitive Mastery process. The model's accurate prediction of test performance on the individual rules both serves to validate the underlying cognitive model in detail and can help guide iterative tutor improvements by identifying those cognitive rules that give students the most difficulty.

Note that students' actual test accuracy is at least 80% for 11 of the 14 rules, and falls to between 60% and 80% for the other three rules. These latter three rules involve reasoning about the expression and transmission of X-Linked recessive traits by males:

- If a trait is X-Linked and recessive and a male is unaffected, Then he cannot be a carrier (because a hemizygous male necessarily displays the trait).
- If a trait is X-Linked and recessive and a daughter has an unaffected father, Then she cannot be a carrier (because her father cannot be a carrier).

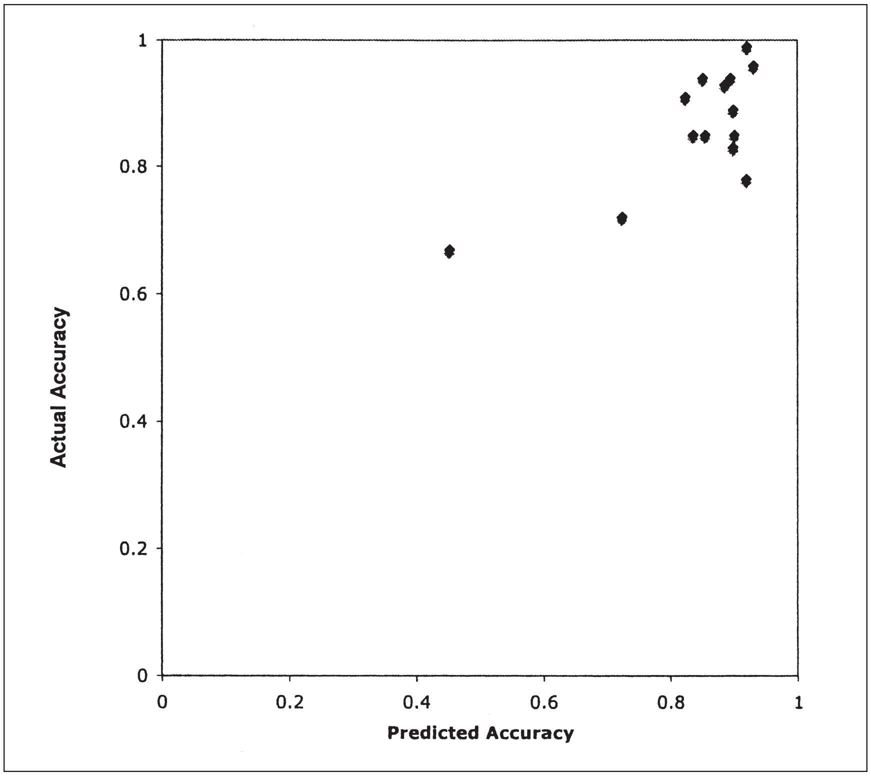

Figure 4. Actual test accuracy plotted as a function of predicted accuracy for 14 rules in the cognitive model.

- If a trait is X-Linked and recessive and an unaffected female has an affected son, Then she must be a carrier (because a son cannot inherit an X-Linked trait from his father).

The Knowledge Tracing algorithm predicts the test accuracy well, suggesting that these three rules are psychologically valid, but the empirical accuracy results can lead us to redesign the Cognitive Tutor curriculum to include more opportunities to apply the rules.

CONCLUSION AND FUTURE PLANS

In summary, in 36 evaluations at 12 institutions, the Cognitive Tutor's Model Tracing support led to sizeable pretest-posttest learning gains for complex multi-step genetics problem solving. These learning gains averaged 18 percentage points or about two letter grades across the evaluations. Detailed empirical evaluations of Knowledge Tracing for one set of students and one module revealed that the student modeling algorithm is capable of accurately predicting individual student performance on paper-and-pencil posttests, and of accurately predicting how well students will apply the individual reasoning rules in the cognitive model.

One important conclusion that emerged is that Cognitive Tutor technology can be successfully applied to genetics problem solving. While successful Cognitive Tutors have previously been developed for middle school mathematics, high school mathematics, and introductory programming, in this project the technology is applied in the substantially different domain of scientific inference. In this domain, students need to reason from observable data to underlying processes as well as to design experiments that will yield useful data. The problem-solving interfaces and underlying cognitive model for the 16 Cognitive Tutor modules proved successful in helping students acquire these reasoning skills.

Designing the problem-solving activities and developing the 500-rule cognitive model for the 125-problem curriculum with its roughly 3000 individual problem-solving steps is a large task. This project has reinforced our prior strong belief that a successful educational technology development process requires an interdisciplinary team, including experts in the problem-solving domain and in teaching in the domain, cognitive scientists with expertise in analyzing and modeling student problem solving and behavioral scientists with expertise in empirical evaluation. Finally, the teacher in the classroom is the other major member of the team in a successful educational technology project. The interest in the project of genetics instructors at our partner sites confirmed our belief that there is a perceived need to provide better genetics problem-solving support across a wide range of colleges and universities. At our summer workshops the instructors were enthusiastic about discussing the challenges they encountered in their courses and provided useful ideas for, and critiques of, the Cognitive Tutor activities. Finally, the instructors were motivated not just to incorporate novel activities into their courses, but to participate in evaluating student learning outcomes. Technology was not a limiting factor in pursuing the project. While instructors had different levels of control over computer facilities at their respective institutions and had different levels of technical support that made software installation more or less challenging, in the end technology did not limit any instructor's participation in the project and the Cognitive Tutor activities proved effective across the various institutions.

Finally, while the project has been successful, we are pursuing multiple activities to further develop the Cognitive Tutor curriculum. In particular, we are developing Knowledge Tracing and implementing Cognitive Mastery for many of the individual modules and we are developing Cognitive Tutor units for additional key topics that arise in undergraduate genetics course, including compound probabilities in inheriting two segregating genes, and human pedigree genetic mapping with log odds analysis.

REFERENCES

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for science literacy.* New York: Oxford University Press.

Anderson, J. R. (1993). *Rules of the mind*. Mahwah, NJ: Lawrence Erlbaum.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4*, 167-207.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Bloom, B. S. (1984). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*, 4-16.

Cavallo, A. M. L. (1996). Meaningful learning, reasoning ability, and students' understanding and problem solving of topics in genetics. *Journal of Research in Science Teaching, 33,* 625-656.

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*, 237-248.

Corbett, A. T. (2001). Cognitive computer tutors: Solving the two-sigma problem. In M. Bauer, P. Gmytrasiewicz, & J. Vassileva (Eds.), *UM2001, User Modeling: Proceedings of the Eighth International Conference* (pp. 137-147). New York: Springer.

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*, 253-278.

Corbett, A. T., & Anderson, J. R. (2001). Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In J. Jacko, A. Sears, M. Beaudouin-Lafon, & R. Jacob (Eds.), *Proceedings of ACM CHI'2001 Conference on Human Factors in Computing Systems* (pp. 245-252). New York: The Association for Computing Machinery.

Garton, G. L. (1992). Teaching genetics in the high school classroom. In M. U. Smith & P. E. Simmons (Eds.), *Teaching genetics: Recommendations and research proceedings of a national conference* (pp. 20-23). Cambridge, MA: National Science Foundation.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8,* 30-43.

Koedinger, K. R., & Corbett, A. T. (2006). Cognitive tutors: Technology bringing learning sciences to the classroom. In R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61-77). New York: Cambridge University Press.

Kulik, J. A. (1994). Meta-analytic studies of findings on computer-based instruction. In E. Baker & H. O'Neil (Eds.), *Technology assessment in education and training* (pp. 9-33). Mahwah, NJ: Lawrence Erlbaum.

Lewis, J., & Wood-Robinson, C. (2000). Genes, chromosomes, cell division and inheritance—Do students see any relationship? *International Journal of Science Education, 22,* 177-195.

National Research Council (NRC). (1996). *National science education standards*. Washington, DC: National Academy Press.

Papatheodorou, I., Kakas, A., & Sergot, M. (2005). Inference of gene relations from microarray data by abduction. In C. Baral, G. Greco, N. Leone, & G. Terracina (Eds.), *Proceedings of the 8th International Conference on Logic Programming and Nonmonotonic Reasoning* (pp. 389-393). New York: Springer.

Smith, M. U. (1988). Successful and unsuccessful problem solving in classical genetic pedigrees. *Journal of Research in Science Teaching, 25*, 411-433.

Zupan, B., Bratko, I., Demšar, J., Juvan, P., Curk, T., Borštnik, U., et al. (2003). GenePath: A system for inference of genetic networks and proposal of genetic experiments. *Artificial Intelligence in Medicine, 29*, 107-130.

Direct reprint requests to:

Dr. Albert Corbett
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA 15213
e-mail: corbett@cmu.edu