Understanding Learning Strategy as Conversation

Steve Battle [1]

Abstract. The goal of this research is to understand the learning strategies employed by students using so called Intelligent Tutoring Systems (ITS), with the aim of supporting efforts to improve the effectiveness of this type of educational technology. An experiment is conducted in which we observe and record online interactions between the participant and the ITS, with the goal of discovering learning strategies used by students. Gordon Pask's *Conversation Theory* defines a social constructivist account of learning and teaching, and the hypothesis explored here is whether analysis of the observed data encoded using CT will help to identify behaviours that represent meaningful learning strategies. These strategies are identified with Markov Processes, enabling the use of *Hidden Markov Models* to discover clusters of distinct patterns of behaviour within the observational data.

1 INTRODUCTION

Intelligent Tutoring Systems (ITS) are teaching software designed to take a student through a series of exercises; facilitating *learning-by-doing* [38]. These exercises are carried out in a simulated environment supporting active learning. We investigate the use of an ITS known as SQLTutor that teaches students 'Structured Query Language' for databases by taking them through a series of practical exercises.

Our goal is to model the *conversations* between the learner and the ITS to identify recurring patterns of interaction with specific *learning strategies*. The foundation of this research is the work of Gordon Pask on *Conversation Theory* [25]. This is a *social constructivist* theory of learning [41, 3] in which the role of the teacher, and the conversations between teacher and student, are made explicit. According to Pask [26], "Learning depends on the strategies used by a student in order to direct his [or her] attention" and "at the other extreme, the strategies may be imposed upon the student as *teaching strategies* by a programme, a teaching device, or a training routine." The term *learning strategy* used in this paper acknowledges this dialectic between teaching and learning strategy in the observed data.

We are particularly interested in the way that a learning strategy may switch between different levels of learning. Pask's *Conversation Theory* [27] states that a conversation can exist on many levels [34], with the lowest level engaged directly with a learning environment. In Conversation Theory, learning becomes a process of negotiation with peers and teachers to construct "shared meaning" [36]. We can relate this to Bloom's Taxonomy of Learning Domains [4, 40] in which there are six major categories of *cognitive* learning processes including simple *recall*, *understanding* and *application* of knowledge, upon which higher level processing involving *analysis*, *evaluation* and ultimately *creativity* are built. These higher level process of comprehension require an awareness of context and the place of this knowledge in the overall *curriculum* [15].

For the purposes of this research, the observational data [6, p.82] may be thought of as a stochastic process [9]. A learning strategy can then be identified as a Markov Process [2] enabling the use of *Hidden Markov Models* (HMMs) to "automatically discover which strategies are used in practice." [5] Furthermore, recognising that learners may use different strategies according to the context and tasks they face, HMM *Clustering* is used to identify and cluster different subsequences into a number of distinct learning strategies [37, 39]. It is hoped that future development of this theory may support a better understanding of how students learn [12, p.9].

2 CONVERSATION THEORY

"Language does not transmit information" claims Humberto Maturana [23, p.57], "its functional role is the creation of a cooperative domain of interactions between speakers." Shannon's Information Theory [35] tells only part of the story, quantifying the information capacity of the channel alone, whereas for learning we are interested in the level of order in the system as a whole [33], comprising both the learner and the teacher. The theory of *cognitive constructivism* rejects the simplistic *transmission* model of information in favour of a model pioneered by Piaget [29, p.77] where, "knowledge results from continuous construction."

Building on Piaget, David Kolb states [17, p.49], "Learning is the process whereby knowledge is created through the transformation of experience." Kolb's Experiential Learning Cycle in Figure 1 is a four-stage cycle describing an effective learning experience. It begins i) with a new or revisited *concrete experience*, followed by ii) observation of, and reflection upon the experience leading to iii) new concept formation, which may in turn iv) suggest new hypotheses which can then be tested in the concrete environment.

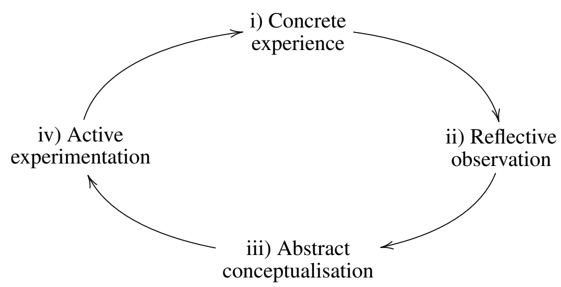


Figure 1: The Kolb Experiential Learning Cycle.

[1] The University of the West of England, Bristol, UK, email: steve.battle@uwe.ac.uk

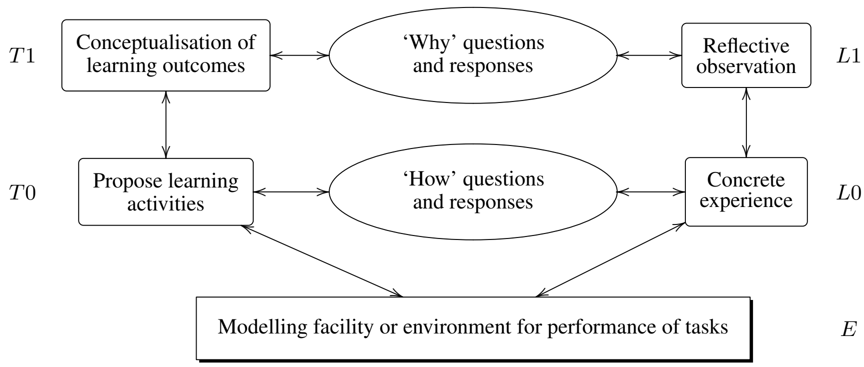


Figure 2: The Skeleton of a Conversation

However, the Kolb cycle is open to criticism [16], the first being that the cycle neglects social learning, particularly the conversation between teacher and learner. The Kolb cycle focuses on the learner as though they were working in isolation. Furthermore, the stages may, in reality, occur out of sequence or not at all. For example, the learner may perform any number of concrete experiments and observations, before later reflecting on the experience as a whole. This would be typical of a practical session where a series of experiments might be conducted en bloc. The reflective stage might be skipped altogether if nothing unexpected occurred that would require the underlying conceptualisation to be modified. Indeed, a learner engaged purely in *surface* learning may simply decide to skip the reflective stage for other reasons; with the reflective and concept formation phases seen as *deep* learning [21, 22, 32].

The theory of social constructivism, developed by Lev Vygotsky [42] rejects the idea of transmission but embraces that of *conversation*. It argues that learning takes place because of a learner's interactions within a group. It supports the idea of reaching understanding by consensus, and it is this idea that finally makes sense of the role of the teacher, and of peers in a learning community. Furthermore, Vygotsky rejected Piaget's notion that a child's *egocentric* language disappears altogether, countering that this private speech is simply a stage in the development of inner speech; an *internal* conversation in all major respects [41].

The cause of Social Constructivism was taken up by British Cybernetician, Gordon Pask, who saw intelligence as an emergent property of *conversation*. From the mid-1950s onward, Pask developed a wide range of teaching machines, the first of which, the Solartron Adaptive Keyboard Instructor (SAKI), led eventually to the development of the world-famous Mavis Beacon typing trainer [30]. His research focused on understanding conversations between teacher and learner, learning peers, a teaching machine interacting with a learner, or in the extreme as an internal dialogue in the mind of an autodidact. Because he is interested in adaptation at a psychological or conceptual level, Pask's *Conversation Theory* [25] is very applicable to these learning conversations. We will use Pask's theory as a framework for understanding the conversation between the learner and an Intelligent Tutoring System in the role of teacher.

Sheila Harri-Augstein and Laurie Thomas, for The Centre for the Study of Human Learning (CSHL) at Brunel University, describe a *learning conversation* [14] as means of recalling experiential memories into awareness, thus being a key enabler for reflection. Moreover, they see learning itself as a skill which can be learned. However, unlike Kolb's Learning Styles and *Experiential Learning cycle*, they view this as a set of parallel conversations occurring at different levels, with task-level conversations ranging over practical activities, and higher level conversations focusing on the learning process itself. We can relate this stratification to Bloom's taxonomy with the task-focused conversation spanning simple recall, through to comprehension involving an appreciation of the wider context, and the place of this knowledge in the overall teaching curriculum. They also observed that traditional teaching practices based on an expert *communicating* their knowledge, may inadvertently reduce a learner's capacity for self-organised learning. This result emphasises the need for good *conversation design* [20] in the development of Intelligent Tutoring Systems.

In his *Conversation Theory*, Gordon Pask develops a similarly stratified model, but one in which the participants of the conversations are explicitly identified. Figure 2 illustrates what Pask calls the 'Skeleton of a Conversation'. The rounded rectangular boxes represent 'Psychological Individuals', or what Pask calls *P-Individuals*. Pask is not particularly concerned with specific learning conversations nor the number of levels, so with Scott [34] we adopt a model that includes individuals that can be loosely identified with Kolb's stages. Both 'Concrete experience' and 'Propose learning activities' deal with the problem of operational learning and involve reasoning directly about action. The 'Conceptualisation of learning outcomes' and 'Reflective observation' work at a higher level, and are concerned with overall comprehension; and are placed correspondingly higher up in the diagram.

The ellipses represent verbal exchanges, or *conversational domains* that psychological individuals are able to engage in. Only individuals at the same level, connected horizontally, interact with each other in this way. The high-level conversations deal with questions about 'Why' the world works the way it does, while the operational levels deal in questions about 'How' the world works; each level having it's own distinctive language.

The vertical connections represent non-verbal, causal connections with feedback. The behaviour of the lower-level process can be reflected upon and selectively modified in the light of some kind of failure; for example, when some action fails to produce the predicted outcome. To ground the operational conversation, individuals at the lowest level level interact with a real or simulated environment; indicated by the heavier box at the bottom of Figure 2. Interaction with this environment is similarly understood as a causal or physical interaction.

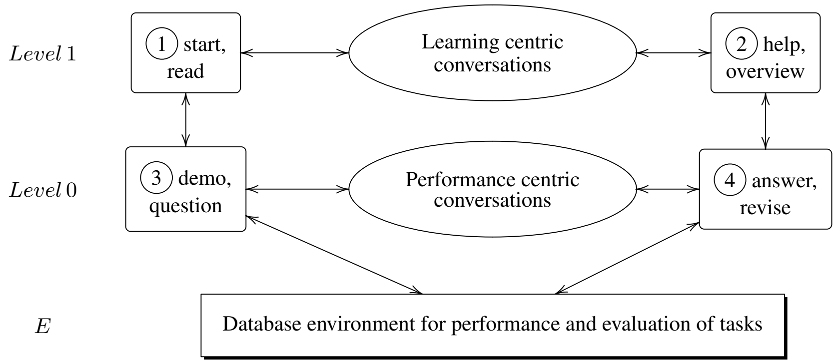


Figure 3: Conversational framework used to encode the data for training the Hidden Markov Models. Each conversational act belongs to one of four classes discriminated by the level of conversation and teaching vs. learning strategy axes.

The P-Individuals in Figure 2 are labelled $T0$, $T1$, $L0$, $L1$ to reflect the most obvious configuration in which the left hand side represents the 'Teacher' ($T0$ & $T1$) and the right-hand side, a 'Learner' ($L0$ & $L1$). Where Kolb describes a solitary learner, this configuration allows us to treat $T1$ as being concerned with the Intended Learning Outcomes whereas $L1$ represents the learners own internally constructed and tacit model of their experience. The Teacher may observe the learner's progress and at $T0$, ask the learner leading operational questions or suggest further experiments. The learner, $L0$, puts these suggestions into practice through interaction with the learning environment E. In this case the processes $T0$ and $T1$ are grouped together as are $L0$ and $L1$. Pask called these groupings *M-Individuals*, or 'Mechanical Individuals', named in a bold attempt to capture both biological and non-biological embodiments of P-Individuals, in a move that allows us to admit teaching machines into the role of Teacher.

Following Kolb more literally, we may arrange individuals $T0$, $T1$, $L0$, $L1$ within a single group – within the mind of an autodidact or self-learner. The self-learner internalises both the teacher and learner, and learning proceeds through an internal dialogue between them. Execution of the cycle would have $L0$ engaging with the environment E; a performance error at this point would be fed up a level to $L1$ for reflection, and then across to $T1$ with the self-learner asking themselves *why* the error occurred. Modifications to the conceptual model may suggest new experiments in the environment E at $T0$ which are then proposed to $L0$, and the cycle repeats.

Conversation Theory also allows us to describe scenarios that are more student-centric, focusing on peer-to-peer collaborative conversations. Collaborative tools are suited to extending the conversation beyond the lecture theatre and have a place in supporting teams of students working together on projects. These are generally conversations at the conceptual level, with $T1$ and $L1$ understood as learning peers; with the model admitting any number of participants into the conversation.

Collaborative learning environments define define a space where learning opportunities are provided in the form of technological artefacts and computer-mediated environments. The so-called flipped classroom [11] focuses on an environment E containing *tangible* artifacts, and *ponderables* introduced by the teacher as facilitator $T0$. Drawing directly on constructivist theories, it focuses on the learning experience of the learner $L0$ and their peers where "the involvement of others can become both catalyst and catharsis in social learning contexts." [13] As before the teacher has a defined set of Learning Outcomes defined at $T1$, and there is an expectation that based on their studies outside the classroom the learner brings Learning Incomes [31] to the table at $L1$.

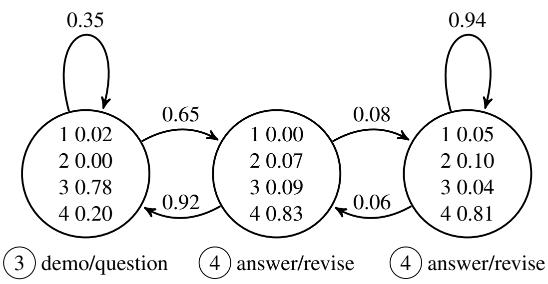
Pask was also concerned with the content of conversations; not the fine-grained structure of knowledge, but the relationships between topics [27]. He describes just two relationships:- *entailment* and *analogical* relationships. A so-called *entailment mesh* represents the important precursors to a given concept [28] and in the ITS this corresponds to the order in which exercises are presented. The sequence of the questions is set by the conversation designer, each question building on previous answers. This idea can be related to the idea of *threshold concepts* [24] acknowledging the necessity of building concepts hierarchically. The ITS can also present examples and ask the student to apply the same idea elsewhere, reinforcing a learned concept through analogy. The resulting heterarchy of topics is more like a web of concepts that can be approached from any angle, and in this case the order is not critical. In the SQLTutor this concept map is represented very simply as a hypertext, and the conversation is the exploration of this curriculum by the learner.

3 EXPERIMENT

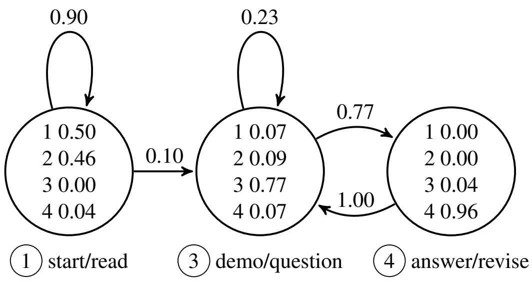
The primary data for this research are collected through observation and recording of online interactions between the participant and the SQLTutor running in a standard web-browser. SQLTutor guides the student through a number of exercises using a sequence of readings, demonstrations, and questions. These questions must be answered in the form of SQL statements that are evaluated by the software and checked against a model answer. It offers guidance where the student's answers are incorrect. While the student is guided along a path that presents challenges in a logical order, the student is actually free to move backwards in the sequence to revise content, or to break out of the sequence entirely and visit any point in the curriculum; no page is more than 3 clicks away from any other. The student may also consult online help or refer to the data model underlying the worked examples. This observational sequence is recorded in a dataset that is uniquely associated with each participant.

Interactions between the participant and the SQLTutor are captured by a proxy server to which the participant directs their browser for the duration of the session. This provides an electronic record of the Uniform Resource Locators (URLs) visited during the experimental session. A short checklist is presented to the participant before the session, to identify any worksheets they may have previously completed with SQLTutor to estimate their prior knowledge. This information, combined with the observed record is used to determine whether participants are revising existing content or exploring new content.

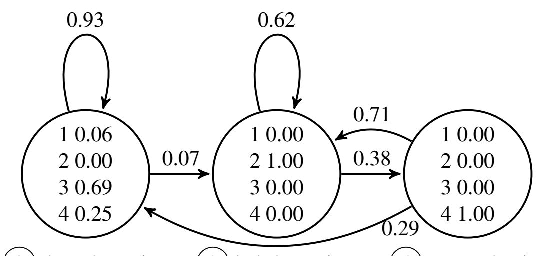
Analysis identifies a number of learning strategies from the observational data. The qualitative intent behind these discovered strategies is then recovered by *triangulating* these observations with follow-up interviews [7, 6]. These semi-structured online interviews, carried out via email [8, p.197], mirror back to those participants relevant excerpts from their observed sessions, asking them what they were trying to achieve. No interpretation was suggested to the participant, leaving them free to provide their own interpretation.



(a) Cluster 1 - *"straightforward Q&A" then "trial and error."*



(b) Cluster 2 - *"seeking help" then "straightforward Q&A".*



(c) Cluster 3 - *"stuck on a question" and "seeking help".*

Figure 5: Three discovered HMM clusters, each one representing a distinct learning strategy. Each state is represented by a node, and transitions between states are shown as directed arcs, labeled with the conditional probability of that transition occurring. In each state a symbol can be emitted, and the probabilities of emitting a symbol (1-4) shown within each state. The symbol with the highest probability is indicated beneath each state along with its conversational acts.

Five subjects participated in the investigation. For analysis these five datasets are merged and segmented at the 'start' points that indicate entry to a new worksheet. Segmenting the data at equivalent points, rather than arbitrarily, allows each HMM to start in the same state. Each visited URL is a page generated by SQLTutor. The recorded URLs are assigned an initial encoding that indicates what kind of page has been visited; the meaning of these *conversational acts* is defined in Table 1. The merged dataset contains 50 sequences, with an average length of 17.28 acts, which are then randomly shuffled to minimise presentation order bias. During training of the HMMs, the shuffled sequences are divided equally into separate training and test sets of 25 sequences each, to reduce over-fitting.

Act	Description
overview	Index listing worksheets, and subjects
start	Worksheet introduction listing contents and related worksheets
read	Topical commentary or summary
help	Help pages and and documentation for the worked example
demo	Demonstration of an SQL command
question	Concrete task proposed by the ITS
answer	Response in the form of an SQL command, and its evaluation
revise	Student revisits an earlier demo, question, or answer page

Table 1: Conversational acts

After Laurillard [19] we re-classify the acts along two dimensions giving us four separate categories as seen in Figure 3. The first (vertical) dimension distinguishes (level 0) task focused or performance centric conversations from higher-level (level 1) conversations which have learning itself as their subject. The learner who lacks the necessary resources to answer a question may explore the range of topics available by returning to a general *overview* page from which they can dip back into and revise specific subject areas. They may also consult general *help* pages accessible from SQLTutor. Performance is grounded in the learning environment, in this case a relational database against which SQL commands can be evaluated, so performance level conversations are those that *demo* a command, ask *questions* & proffer *answers*, or *revise* earlier tasks. Along the second (horizontal) dimension we discriminate between actions initiated by the SQLTutor, such as the *start* of a workshop or an introductory *read*ing, from those initiated by the student learner. This dimension throws into relief the differences between teaching and learning strategies.

This provides us with a reduced set of four encodings, representing four distinct classes of conversational act within Conversation Theory. This merging of conceptually similar acts significantly reduces the size of the search space for Hidden Markov Model clustering.

1. **start/read** (teacher initiated, learning centric)
2. **help/overview** (learner initiated, learning centric)
3. **demo/question** (teacher initiated, performance centric)
4. **answer/revise** (learner initiated, performance centric)

4 HIDDEN MARKOV MODEL CLUSTERING

Hidden Markov Models are used to model potential learning strategies. HMMs are used because we are only able to observe the external actions that appear in a conversation. We do not, of course, have access to the internal 'mental' states of the learner. A HMM is a graphical model where each state is represented by a node, and transitions between states are shown as directed arcs, labeled with the conditional probability of that transition occurring. In each state

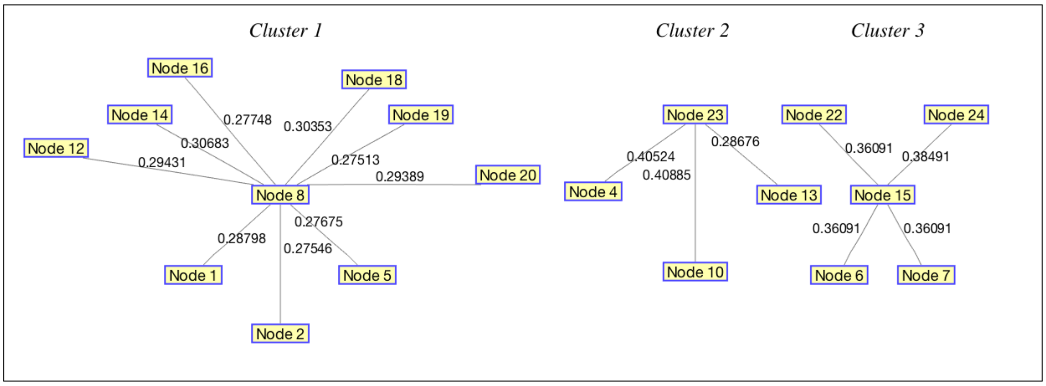


Figure 4: Thresholded Minimum Spanning Tree used to recover clusters. Nodes 1 to 25 represent the segments of observational data from the test set. Another 25 segments comprise the training set. Singleton nodes are not shown.

a symbol can be emitted, here representing one of the four conversational classes of action. The *Markov property* requires that the state at time $t + 1$ is *only* dependent on the state at time t; nothing that happened prior to that is relevant. Furthermore, the probability of observing a symbol at time t depends only on the state at time t. If there are any historical dependencies, these must be somehow encapsulated in the current state, in other words the current state is the only 'memory' of the HMM.

Each HMM is defined by a pair $\langle T, E\rangle$ defining a transition matrix T and an emission matrix E. These matrices hold the conditional probability of a given transition, and emission of a symbol, given the current state. The HMMs representing distinct strategies are initially unknown so an HMM clustering algorithm [37] is employed to learn the parameters for a set of HMMs that are, in effect, competing to represent the observed data.

The MATLAB® Statistics and Machine Learning Toolbox[2] is used to train a set of HMMs to recover the (hidden) sequence of states that generated a given set of observed data. The clustering algorithm takes an initially random set of HMMs and partitions the data by finding the best fitting HMM for each datum (a sequence). The number of states and clusters is deliberately kept low (between 1 and 3) to counter over-fitting to the data. The HMM is then trained on that partition (5000 iterations) and the data is re-partitioned and the process repeated for a fixed number of iterations (50). This employs the MATLAB function *hmmtrain* using the Baum-Welch algorithm [1] which is a hill-climbing search that will converge to a local maximum from an initial randomised HMM.

To achieve a global maximum an evidence accumulation technique is used [10] in which co-occurrences of pairs of sequences in the same cluster count as votes towards their eventual association. In addition these co-occurrence counts are weighted by the quality of the solution as evaluated against the test set. After many voting rounds (5000), a consensus global clustering is achieved.

To recover the clusters from the weighted co-occurrence matrix, a minimum-spanning-tree is constructed [18] over the nodes representing the input sequences and the weakest links are cut below a threshold in order to form the clusters shown in Figure 4. The individual nodes in this figure are sequences of observational data from the test set, and it can be seen that there are three clusters, labeled as clusters 1 to 3.

Representative HMMs are then regenerated from the resulting clusters, again using MATLAB's *hmmtrain* function, as shown in Figure 5. Each has 3 states with the transition probabilities indicated on the directed arcs, and the probabilities of emitting a symbol (1-4) (the emission vector) is shown within each state. These symbols correspond to the classes in the framework of Figure 3, and to assist interpretation the symbol with the highest probability is shown beneath each state along with its corresponding conversational acts.

In most cases the HMM training has identified a single high-probability emitted symbol for each state. The only exception to this is the first state in cluster 2 with two approximately equiprobable symbols (1 and 2). This means that the strategy represented by this state does not distinguish between Level 1 learning-centric acts, but does distinguish these from Level-0 performance-centric acts.

4.1 Triangulation by thematic analysis

Given the results of this analysis, the next step was to conduct semi-structured interviews with the participants to obtain their views on the strategies they were using at the time. This enables us to *triangulate* the results of the observational analysis with qualitative comments from each participant, who were asked about representative examples of their own activity falling within different clusters. Not all participants exhibited activity from every cluster. Table 2 summarises the results of these interviews. A simple *thematic analysis* of the comments, generalising across participants, allows us to pull out qualitative features which we can use to describe each cluster.

An NVivo®[3] based analysis enables us to encode the appearance of specific keywords within a comment. These keywords are aggregated to code for certain terms suggested by the discovered learning strategies. These terms may then be combined in ways that reflect the learning strategies more closely. For example, the combination of terms, "seeking help" and "straightforward Q&A" should select for strategies matching cluster 2.

To test the hypothesis that these combinations of terms discriminate between the Hidden Markov Model clusters a contingency table is constructed in Table 3. The rows and columns of this table are nominal values; clusters 1 to 3 are arranged in columns, and each row represents criteria expressed as a term or combination of terms. A row called 'other' captures any unmatched comments, so that all comments are accounted for. The *null* hypothesis is that these criteria and the the clusters are independent. A chi-square test cannot be used

[2] MathWorks http://uk.mathworks.com/products/statistics/

[3] NVivo http://www.qsrinternational.com/nvivo-product

here because the expected values are less than 5, so an *exact contingency table test* is used to calculate the probability of the observed table. The sum of probabilities for this table is $p = 0.008 < 0.05$ (5% significance), so we reject the null hypothesis and conclude that the terms found in thematic analysis are sufficiently discriminatory. This provides additional support for the learning strategies discovered by Hidden Markov Model clustering.

The following points enumerate the terms used to encode the participant's comments, represented as aggregations of keywords.

- **trial & error**
  - "wrong"
  - "trial and error"
  - "see what would happen"
  - "attempts", "attempting"
- **stuck on a question**
  - "stuck"
- **straightforward Q&A**
  - "valid"
  - "trying"
  - "straightforward"
  - "confident"
- **seeking help**
  - "structure of the tables"
  - "look"
  - "help"
  - "familiarise"

Drawing on this thematic analysis we are now in a position to construct an interpretation of the Hidden Markov Models.

Cluster 1 - *"straightforward Q&A" then "trial and error."* in Figure 5(a) begins with a cycle of "straightforward Q&A", of demo/question then answer/revise. There is repetition around 'demo/question' where SQLTutor provides a sequence of demonstrations. However, a difficult question may be followed by a flurry of "trial & error" activity where the same question is answered incorrectly multiple times. Crucially, and for whatever reason, the student is not accessing the help system. According to Sleeman [38] incorrect answers are not always bad, "some floundering can be vitally important. The crucial meta-skill of knowing when one's floundering is useless can only be discovered by trial and error." In other words, this floundering can be seen as a learning strategy in which the learner comes to understand how to apply their knowledge.

Cluster 2 - *"seeking help" then "straightforward Q&A".* in Figure 5(b) is often seen at the beginning of a session where the learner is following the provided reading matter while trying to familiarise themselves with the learning environment. This *serial* activity [26, p.269] around reading is expected as this is the default offered by the available teaching strategy. In the discovered learning strategy this is seen to be followed by a sequence of "confident" question answering where the subject correctly answers all the questions. This combination is most likely to be seen early on when the subject is both unfamiliar with the environment and the questions are simpler.

Cluster 3 - *"stuck on a question" and "seeking help".* in Figure 5(c) contains a question and answer with multiple visits to help pages interposed between them. Even when an (incorrect?) answer has been given, there is high probability of need for further help.

Cluster	Subject	Comment
1	1	"...simply me working through the worksheet normally."
1	1	"...got myself confused using the program, was attempting to recall how to access certain parts."
1	1	"Stuck on a particular question, came down to syntax of SQL statement."
1	2	"to pull information from multiple tables in a single query."
1	3	"likely to be the straightforward questions, or pages that Ive (sic) opened only once."
1	3	"questions that I got wrong multiple times. I used trial and error to get to the answer."
1	4	"The simple SQL statements sections was fairly straight forward."
1	4	"With this cluster I got the arithmetic wrong, ... after a few attempts though I finally realised my mistake and corrected it."
1	5	"I'd forgotten the 'emp' table name, so I just put in the incomplete SQL fragment to see what would happen."
2	2	"I was trying to familiarise myself with SQLTutor and trying out the SQL sandbox."
2	3	"questions that I was confident about. Got all of those on first try."
2	4	"I started off a little confused and began going back and forth between the help sections and the first task."
3	1	"was stuck on a question, so went to look for information within the program."
3	3	"getting familiar with the software and structure of the tables."
3	5	"I just needed to use the DESCRIBE statement for two tables. I already had the 'help' page open..."

Table 2: Triangulation of the results of the observational analysis with qualitative comments from each participant, who were asked about representative examples of their own activity falling within different clusters. Not all participants exhibited activity from every cluster.

Cluster:	1	2	3
trial & error	4	0	0
*seeking help **and** straightforward Q&A*	0	2	0
*stuck on a question **and** seeking help*	0	0	2
other	5	1	1

Table 3: Contingency table where clusters 1 to 3 are arranged in columns, and each row represents criteria expressed as a term or combination of terms. The discrimination provided by these criteria provides additional support for the learning strategies discovered by Hidden Markov Model clustering.

5 CONCLUSION

This investigation demonstrates the efficacy of using HMM clustering to discover learning strategies in observational records of interactions with an ITS. The benefit of HMMs is that they provide a bottom-up approach to learning strategy discovery, rather than defining them in advance of the investigation. These discovered strategies are seen to recur within sessions and, perhaps more significantly, across multiple students. *Triangulation* enables us to recover student intent via interviews, and verify that these learning strategies have a stable meaning. This confirms the hypothesis that Conversation Theory provides an adequate framework for understanding learning strategies.

ACKNOWLEDGEMENTS

I would like to thank Chris Wallace for the continuing development of SQLTutor[4], used extensively in our teaching of Structured Query Language.

REFERENCES

[1] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, 'A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains', *Annals of Mathematical Statistics*, **41**(1), 164–171, (1970).

[2] Leonard E. Baum and Ted Petrie, 'Statistical inference for probabilistic functions of finite state markov chains', *Ann. Math. Statist.*, **37**(6), 1554–1563, (12 1966).

[3] P.L. Berger and T. Luckmann, *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*, Open Road Media, 2011.

[4] B.S. Bloom, *Taxonomy of Educational Objectives: The Classification of Educational Goals*, number v. 1 in Taxonomy of Educational Objectives: The Classification of Educational Goals, D. McKay, 1956.

[5] Kristy Elizabeth Boyer, Eun Young Ha, Michael D. Wallis, Robert Phillips, Mladen A. Vouk, and James C. Lester, 'Discovering tutorial dialogue strategies with hidden markov models', in *Proceedings of the 2009 Conference on Artificial Intelligence in Education: Building Learning Systems That Care: From Knowledge Representation to Affective Modelling*, pp. 141–148. IOS Press, (2009).

[6] E. Cleaver, M. Lintern, and M. McLinden, *Teaching and Learning in Higher Education: Disciplinary Approaches to Educational Enquiry*, SAGE Publications, 2014.

[7] Louis Cohen, Lawrence Manion, and Keith R. B. Morrison, *Research methods in education (5th edition)*, Routledge, Oxon, UK, 2000.

[8] M. Denscombe, *The Good Research Guide: For Small-scale Social Research Projects*, Open UP study skills, McGraw-Hill Education, 2014.

[9] J.L. Doob, *Stochastic Processes*, Wiley Publications in Statistics, John Wiley & Sons, 1953.

[10] Ana L. N. Fred and Anil K. Jain, 'Combining multiple clusterings using evidence accumulation', *IEEE Trans. Pattern Anal. Mach. Intell.*, **27**(6), 835–850, (June 2005).

[11] P. Freire, *Pedagogy of the Oppressed*, Continuum, 1970.

[12] H. Fry, S. Ketteridge, and S. Marshall, *A Handbook for Teaching and Learning in Higher Education: Enhancing Academic Practice*, Taylor & Francis, 2003.

[13] Susan Greener, 'Flipped or blended? what?s the difference and does it make a difference to learning in he?', in *Proceedings of the 10th International Conference on e-learning (ICEL2015)*, (June 2015).

[14] S. Harri-Augstein and L.F. Thomas, *Learning Conversations: The Self-Organised Learning Way to Personal and Organisational Growth*, Routledge, 1991.

[15] Janet Holland and Marcus Childress, 'conversation theory conceptualized in e-learning environments', in *Handbook of Conversational Design for Instructional Applications*, IGI Global, (2008).

[16] Peter Jarvis, 'Learning to be a person in society', in *Contemporary Theories of Learning: Learning Theorists ... In Their Own Words*, ed., K. Illeris, Taylor & Francis, (2009).

[4] 'The Wallace Line' http://thewallaceline.blogspot.co.uk/2009/02/data-modelling-tutor.html

[17] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*, Prentice-Hall, Englewood Cliffs, New Jersey, 1984.

[18] Joseph B. Kruskal, 'On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem', *Proceedings of the American Mathematical Society*, **7**(1), 48–50, (February 1956).

[19] D. Laurillard, *Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies*, Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies, Routledge/Falmer, 2002.

[20] Rocci Luppicini, 'Introducing conversation design', in *Handbook of Conversational Design for Instructional Applications*, IGI Global, (2008).

[21] F. Marton and R. Säljö, 'On qualitative differences in learning: I – outcome and process', *British Journal of Educational Psychology*, **46**(1), 4–11, (1976).

[22] F. Marton and R. Säljö, 'On qualitative differences in learning: II – outcome as a function of the learner's conception of the task', *British Journal of Educational Psychology*, **46**(1), 115–127, (1976).

[23] Humberto R. Maturana and Francisco J. Varela, *Autopoiesis: The Organization of the Living*, Springer Science & Business Media, 1993.

[24] J. Meyer and R. Land, *Overcoming Barriers to Student Understanding: Threshold Concepts and Troublesome Knowledge*, Taylor & Francis, 2006.

[25] G. Pask, *Conversation, Cognition and Learning: A Cybernetic Theory and Methodology*, Elsevier, New York, 1975.

[26] G. Pask, *The Cybernetics of Human Learning and Performance: A Guide to Theory and Research*, Hutchinson Educational, Hutchinson Educational, 1975.

[27] G. Pask, *Conversation Theory: Applications in Education and Epistemology*, Elsevier, 1976.

[28] G. Pask and B.C.E. Scott, 'Caste: A system for exhibiting learning strategies and regulating uncertainties', *International Journal of Man-Machine Studies*, **5**(1), 17 – 52, (1973).

[29] J. Piaget, *Genetic epistemology*, Columbia University Press, New York, 1970.

[30] Andrew Pickering, *The Cybernetic Brain: Sketches of Another Future*, The University of Chicago Press, Chicago, 2009.

[31] P. Race, *The Lecturer's Toolkit: A practical guide to assessment, learning and teaching*, Taylor & Francis, 2014.

[32] P Ramsden, *Learning to teach in higher education*, Routledge, 2nd edn., 2003.

[33] Erwin Schrödinger, *What is Life?*, Cambridge University Press, Cambridge, 1944.

[34] Bernard Scott, 'Conversation theory: A constructivist, dialogical approach to educational technology', *Cybernetics and Human Knowing*, **8**(4), 25–46, (2001).

[35] Claude E. Shannon, 'A Mathematical Theory of Communication', *The Bell System Technical Journal*, **27**(3), 379–423, (1948).

[36] Mike Sharples, 'Learning as conversation: Transforming education in the mobile age', in *Seeing, Understanding, Learning in the Mobile Age*, pp. 28–30, (2005).

[37] Benjamin Shih, Kenneth R Koedinger, and Richard Scheines, 'Discovery of student strategies using hidden markov model clustering', *International Conference on Educational Data Mining*, (2010).

[38] D. Sleeman and J.S. Brown, *Intelligent tutoring systems*, Computers and people series, Academic Press, 1982.

[39] Caitlin Tenison and Christopher J. MacLellan, 'Modeling strategy use in an intelligent tutoring system: Implications for strategic flexibility', in *Intelligent Tutoring Systems: 12th International Conference, ITS 2014, Honolulu, HI, USA, June 5-9, 2014. Proceedings*, eds., Stefan Trausan-Matu, Kristy Elizabeth Boyer, Martha Crosby, and Kitty Panourgia, 466–475, Springer International Publishing, (2014).

[40] Kenneth E. Vogler, 'Improve your verbal questioning', *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, **79**(2), 98–103, (2005).

[41] Lev Vygotsky, *Thought and Language*, MIT Press, Cambridge, MA, 1962.

[42] Lev Vygotsky, *Mind in Society*, Harvard University Press, 1978.