Hidden Markov models
and learning in authentic situations

Léon Harvey
Université du Québec à Rimouski

This paper introduces Hidden Markov Models for the analysis of authentic learning data from an applied field. For illustrative purposes, it shows how classical 2-state all-or-none models can be extended to adequately fit the competence development process of nursery apprentices in a clinical context. It also presents some of the main underlying ideas, such as model specifications, parameters estimation, model selection, the Viterbi algorithm, and goodness-of-fit issues.

Markov models have been used in psychology since the mid-fifties (Miller, 1952; Steiner & Greeno, 1969) to infer cognitive states from sequences of data in learning experiments. They are now considered very general tools for integrating large sets of longitudinal observations (Langeheine, Stern, & van de Pol, 1994), from implicit learning (Visser, Raijmakers, & van der Maas, 2009) to well-being (Eid, & Langeheine, 2007). They have also been used in the classroom context to study negotiations between actors (Weingart, Prietula, Hyder, & Genovese, 1999), peer scaffolding (Pata, Lehtinen, & Sarapuu, 2006) emerging from interactions between students in a synchronous network environment, and to compare counselling methods used by effective and ineffective students (Duys and Headrick, 2004). Some advanced models are also developed to account for sequential decision processes (Fu & Anderson, 2006; Littman, 2010; Niv, 2009).

In this paper, the process of elaborating a Hidden Markov Model (HMM) is presented for tutorial purposes. It reaffirms some of the main ideas underlying anterior tutorial works (Visser, Raijmakers, & Molenaar, 2002; Wickens, 1982) and proposes, for illustrative purposes, to extend the classical 2-state model to observations from a clinical field. Modeling such data from an applied field is an important contribution. It suggests that HMMs are of great practical value when synthesizing competence development processes in authentic learning situations.

More specifically, some HMMs will be built to illustrate how the interactions between a nursery supervisor and her apprentices can be analyzed to grasp the hidden process of competence growth of apprentices in a clinical context. Meanwhile, the tutorial discusses some important issues, such as model specifications, parameter estimation, model selection and goodness-of-fit issues.

Model specifications

Model specifications arise from intertwined psychological and mathematical considerations. Thus, a Markov chain is defined as a series of states. The main property of all Markov chains is that knowledge of an "S" state at time "t" is all that is necessary to predict the evolution of the system at "$t+1$." A conditional probability (1) expresses this idea and should be read as the probability that an individual goes to state S', considering that he was previously in an S_t state, only depends on that S_t state and not on previous ones.

$$P(S_{t+1} = S'/S_1,\dots\ S_t) = P(S_{t+1} = S'/S_t) \qquad (1)$$

A **T** function specifies the transition probabilities $p(S'/S)$ from state S to S'. A π vector defines the initial probabilities $\pi_i = p(S_i)$ of being in S_i states just before starting the observation process. The elements of this vector and the transition function are constrained and equal 1. These can be expressed as $\sum_i \pi_i = 1$ and $\sum_i p_i(S'/S) = 1$.

In psychology, states refer to the internal dispositions of an individual and these cannot be directly observed. For instance, literature in psychology and education largely suggests that the competence of a person cannot be known

by simply looking at her; it must be inferred from her performance during a series of authentic tasks. So, by definition, competences are considered hidden constructs. As a consequence, a special family of Markov models called Hidden Markov Models are recommended to analyze them.

In HMMs, both a set of manifested behaviours and a set of hidden states must be defined. Observation function O then specifies the conditional probability of observing a manifested behaviour (Ω) at each S state. These conditional probabilities $p(\Omega_i \ / \ S)$ are submitted to the constraint $\sum_{i=1} p(\Omega_i/S) = 1$.

In general learning models (Visser et al.., 2002; Wickens, 1982), three categories of states are proposed to account for success and error sequences. These models distinguish the learned, error and intermediary states. The learned states are characterized by the probability of errors approaching zero. In most models, they are considered as absorbing states and this expresses the idea that once a concept or a rule has been mastered, the learner does not revisit the error or intermediary states. The error states are characterized by a probability of errors and successes near random performance. Intermediary states correspond to partial mastery levels and the probabilities of observing a success are higher than expected for a random behaviour.

Moreover, in many learning experiments, the set of observations are composed of two types of results called error and success. However, this set can be extended. For instance, in many professions, important concepts and procedures are learned in the context of authentic interventions as part of the apprenticeship. To account for competence growth, observation data are often gathered using observation grids that record multiple variables about the apprentice and the work situation. These grids record important aspects about the general level of autonomy of the apprentice, whether or not she correctly conceptualizes what she has to do, and whether or not she displays the correct procedural behaviour without error or forgetting. For instance, such grids are especially useful in describing the performance of nurses in an apprenticeship context (Harvey & Barras, 2008; Harvey, 2009).

However, it is necessary to adopt a notation to systematize such data from complex situations. Therefore, to generate a meaningful unit from these observations, bigrams or trigrams can be created. For example, Weingart et al. (1999) have used this strategy to study negotiations. They joined two dichotomous variables to create a bigram that indicates the person in turn-taking and qualifies the nature of the interaction (distributive or integrative). A trigram is composed of three letters, a letter for each of the variables observed. For each letter, an uppercase indicates an adequate scheme, while a lowercase stands for an inappropriate one. Thus, to systematize our notation, the first letter of the trigram will be indicative of the adequacy of the conceptual scheme (*C or c*) based on the explanations given before and during the intervention. The second letter will design the quality of the procedural schemes (*P or p*) associated with the instrumental aspects of the intervention. This letter is set to *p* when errors or forgetting occurs or to *P* otherwise. Finally, the third letter is an overall judgment about the autonomy (*a or n*) of the apprentice. The space of observations is consequently defined as the set of trigrams Ω = {*cp_n, cp_a, Cp_n, Cp_a, cP_n, cP_a, CP_n CP_a*}.

Some of these trigrams have strong psychological meanings. For instance, CP_a represents the ideal situation where the apprentice has all the conceptual and procedural schemes to perform a task and is autonomous. On the other hand, cp_n corresponds to a situation where the apprentice is not autonomous and failed to mobilize adequate knowledge schemes in the situation. Notice that some trigrams are not expected to be observed very often. These are cp_a, CP_n, cP_n and cP_a. First, cp_a represents an hypothetical situation where the apprentice does not possess the adequate conceptual and procedural schemes but would have been considered autonomous. Second, CP_n is the opposite situation where all the appropriate schemes are observed but the apprentice is nevertheless not considered autonomous for some (mainly affective) reasons. Finally, cP_n and cP_a are situations where the procedural schemes would be present but the apprentice would be unable to explain his actions. All of these situations are considered unlikely, at least, in an explicit learning context.[1]

---

[1] Note that the distinction between declarative and procedural schemes made here is not new (Harvey & Anderson, 1996; Singley & Anderson, 1989) but is fundamental in education (Potgieter, Harding, & Engelbrecht, 2008). The notion of schemes refers very largely to what in a situation is transposed to similar situations or generalized across situations. Schemes are defined as a form of abstract mental representation that guides action (Sabah, 2002). They are the basic internal resources used by the competences. Schemes are transformed and attuned from situation to situation. Declarative schemes are general knowledge structures organized into interrelated networks of concepts. They must be distinguished from more context-specific procedural schemes represented by the "condition-action" rule. The basic assumption is that such an interrelated network of concepts is needed by the information processing machinery to create procedures that are appropriate to a given context.
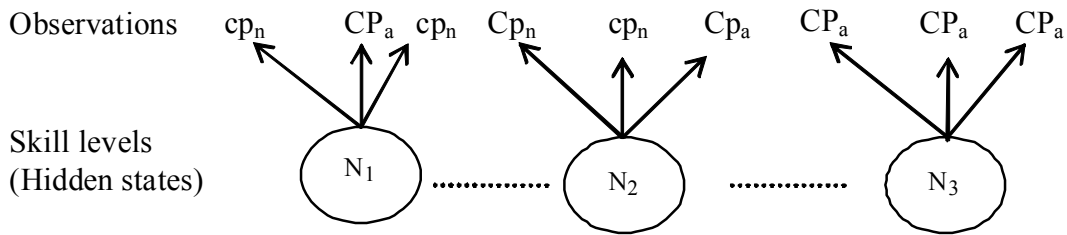
Figure 1. General model. Skill levels *N* (hidden states) are inferred from the observations.

It is possible to intuitively map these observations into hidden competence states. Figure 1 illustrates the mapping process. First, the learned state is expected to be composed mainly of CP_a observations. The unlearned state should be characterized both by the presence of cp_n episodes and the absence of autonomy in more than 50% of the cases. Intermediary states are expected to be composed of mixture of observations with autonomy level greater than 50%.

Moreover, different learning models can be specified to grasp the competence development process. For instance, the literature suggests that some conceptual changes occur as a learner progressively abandons their naïve and intuitive conception of a situation to construct more elaborate explanations (Vosniadou, 2007). Such a change can be symbolized as a transition between two states (symbolized $cp \rightarrow Cp$). Bruner, Goodnow and Austin (1977, p. 50) speak of a "…transition experience between not having a distinction and having it." The literature also identifies many mechanisms operating on procedures and concepts to create meta-procedures (symbolized $cp \rightarrow cP$). The conceptual knowledge base also supports the development of procedural ones through problem solving (symbolised $Cp \rightarrow CP$). On the contrary, it is also accepted that previously learned procedures support the acquisition of new concepts by *a posteriori* reflection on action (symbolised $cP \rightarrow CP$). When some successive transitions occur, paths with 2, 3, or more states might occur. As a result, using this terminology, a 2-state model is a model where a learner transits from a state characterized by the absence of adequate conceptual and procedural schemes to a state where these schemes have been acquired ($cp \rightarrow CP$). Note that this process occurs without visiting some intermediary states. Figure 2a illustrates this 2-state model.

A step model emerges when a learner transits from the unlearned state to the learned state by visiting some intermediary states (Wickens, 1982). Therefore, a 3-state step model has such an intermediary state. Figure 2b illustrates a 3-state step model. Notice that a 1-state model can also be observed. Such a model suggests that the apprentice remains in the same state for the entire period of observation. This may occur if the apprentice fails to learn or, alternatively, if she is already in the learned state at the beginning of the observation process. How to determine the most likely models from the observations will be discussed in the next sections.
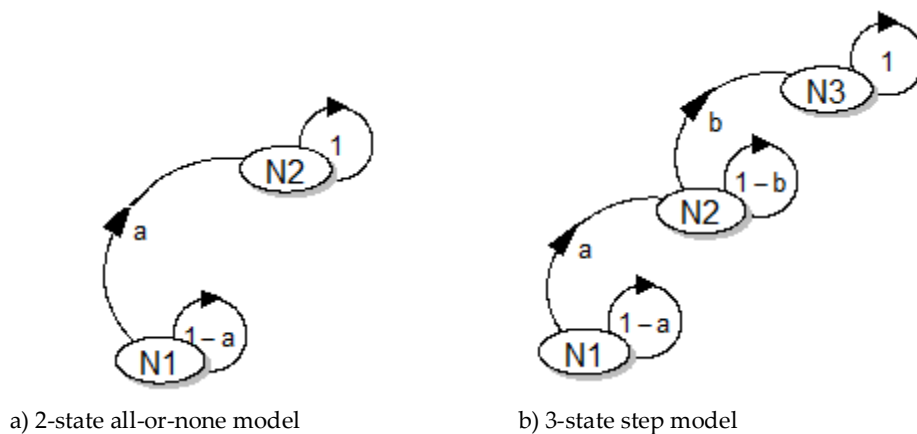

a) 2-state all-or-none model   b) 3-state step model

Figure 2. Some paths for the acquisition of conceptual and procedural schemes during skill learning.

Parameter estimation

Parameter estimates of a given model are obtained iteratively by maximizing the likelihood of the observed sequences using the Expectation/Maximisation (EM) algorithm (Visser et al, 2002). This algorithm searches for optimal parameters that best describe the series. For long series, it is recommended (Hélie, 2006) to maximize the log-likelihood of the data. However, as a result, only local optima might be found and results must be interpreted with caution. It is important to notice that the number of parameters to estimate changes from model to model. For instance, our 2-state model is composed of 22 parameters if all types of observations are made. Indeed, it has 2 parameters for the initial distribution π, four for the transition function, and 16 (2 x 8) for the observation function. However, as π is constrained and sums to 1, only one parameter of the initial distribution is free. Similarly, the transition function has 2 free parameters and the observation function 14. Consequently, the number of free parameters for our 2-state model is 17. Alternatively, our 3-state model might have up to 36 parameters (3, 9, 24) with 29 free ones. These models may have a smaller number of parameters if some observation types are not present in the data.

Model selection and model fit

Once the parameters of a model have been estimated, the next step is to determine how well it fits the data (Hélie, 2006; McCoach & Black, 2008; Visser et al., 2002). Two kinds of questions are of interest (Wickens, 1982). The first is whether a more complicated model is an improvement over a simpler one. The second concerns the overall fits of the model.

Strategies for comparing and selecting models have received much attention in recent years in the literature (Hélie, 2006; McCoach & Black, 2008). First, the significance of each parameter can be tested. If some parameters are not significantly different from zero, a simpler model might be preferred. Second, there are different tests based on the likelihood of the series. For instance, when comparing models with the same number of parameters, the likelihood-ratio test can be used (Hélie, 2006; Visser et al., 2002). This test indicates whether or not two models are significantly different based on the ratio between their likelihoods. When comparing models with a different number of parameters, their likelihoods are not directly comparable and alternative indices, such as the Aikaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC), are competing choices. These indexes make corrections to the likelihood according to the number of degrees of freedom across

models (AIC and BIC) and the length of the series (BIC). They are defined as (Visser et al, 2002, p. 190):

$$AIC = -2 \log L + 2n_{\mathrm{P}} \qquad (2)$$

$$BIC = -2 \log L + n_{\mathrm{p}} \log(N), \qquad (3)$$

where L is the likelihood of the fitted model, n_p the number of free parameters of the model and N, the number of observation used in fitting the model. As a rule of thumb, AIC index generally favors the selection of more complex models while the BIC advantages simpler ones (Hélie, 2006).

The second question of interest is to determine if it adequately fits the observation data. Here, a goodness-of-fit test can provide valuable information. A property of the Markov model is that it can predict the distribution of observations based on the theoretical state at each time period. For this, the Viterbi algorithm provides the optimal hidden states sequence according to the observations at hand and the specified model. When many sequences are analysed with the same model, the Viterbi algorithm can be applied successively to each independent realisation (subject). Then, a Chi-square goodness-of-fit test can be calculated based on observed and predicted observations.

$$\chi^2 = \sum \frac{(F_o - F_E)^2}{F_E}, \qquad (4)$$

Such a test has *c-p-1* degrees of freedom when the estimation of parameters and the goodness-of-fit test are done on the same data set (Visser et al, 2002; Wickens, 1982). Here, *c* is the number of categories used in the test and *p* is the number of free parameters in the model. So, with 169 observation periods and a 2-state model with 17 free parameters, the χ^2 would have 151 degrees of freedom (169 - 17-1).

Skill development during apprenticeship in a clinical context

For illustrative purposes, data from authentic clinical situations will highlight that HMMs can be valuable tools to synthetize the learning process. For instance, in the health care system and nursing in particular, literature points out that some contextual factors may inhibit learning and have a negative impact on the knowledge gained (Lauder, Reynolds & Angus, 1999). In order to characterize the learning process in such a controversial context, this paper proposes that the interactions between a supervisor and 13 apprentices be traced and modeled using HMMs.

As will be shown, the HMM are very useful and informative as they can clearly characterize the learning process of each apprentice (using the Viterbi algorithm). For each situation, a set of variables about the apprentice's cognitive states will be observed. Then 1-state, 2-state and 3-state models will be tested. Although the actions of the
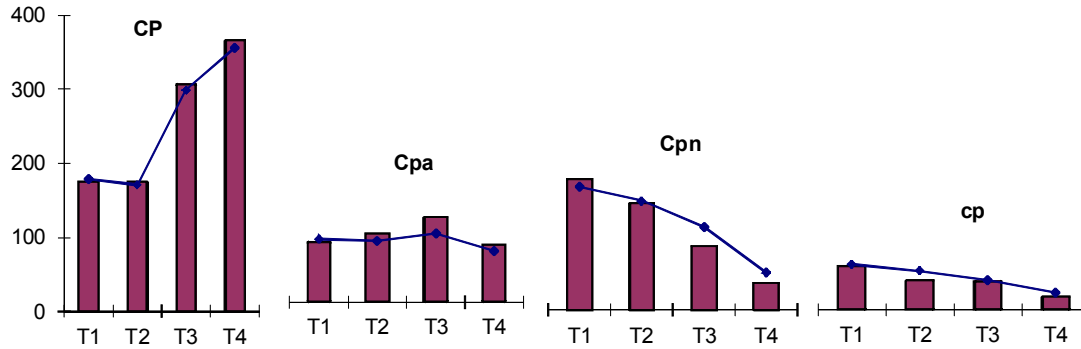
Figure 3. Observed and expected frequency distributions of *CP*, *Cp_a*, *Cp_n* and *cp* data over each day (T1 to T4).

supervisor are also recorded, they will not be integrated into the HMM. Such integration would require a more complex Markov model (e.g. Littman, 2010) and is outside the scope of this paper. The next section briefly reviews the methodology of the study.

The observation set is from Harvey and Barras (2008; Harvey, 2009). Thirteen students enrolled in a nursery curriculum at two colleges participated on a voluntary basis in the study. There were 11 females and 2 males. They were observed during health care activities with 43 patients in two hospitals as parts of their professional training.

Two skills have been observed as defined by the nursery care program in the province of Québec (Ministère de l'Éducation, du Loisir et du Sport, 2004). These skills can be translated as "*Intervene and provide health care to aged people with loss of autonomy in a health care institution*" and "*Intervene and provide health care to adults and aged people in medicine and chirurgical units*." To master these skills, an apprentice must be able to complete a large variety of activities. These activities generally follow a "*plan, act, evaluate and follow-up*" cycle. Each intervention has been further broken down into episodes. Each episode corresponds to an action of the apprentice. In total, 1,926 episodes were observed. A series of episodes forms a sequence. Thirteen sequences have been observed, one for each subject. On average, there are 148.15 episodes per sequence with a standard deviation of 12.56 and a range between [126 and 169] episodes.

An observation grid similar to the one used in the academic institution has been used and modified to better suit the needs of the study. It distinguishes 1) the intervention performed; 2) the quality of the apprentice's explanations (adequate or not); 3) presence or absence of errors or forgetting; 4) the scaffold actions {Demonstration, Coach, Help, Information, Observation}. Moreover, the supervisor has noticed 5) whether or not an appropriate follow-up has been provided and 6) has made a binary

judgment on the autonomy of the apprentice (autonomous versus not autonomous).

The analyses of this data set are divided into two parts. First, trigrams are created from the action of joining three variables. Thus, each episode has been coded as one of the observation sets {*cp_a*, *cp_n*, *Cp_a*, *Cp_n*, *cP_a*, *cP_n*, *CP_a*, *CP_n*} by joining the *Explanations* (*C or c*), *Errors and Forgetting* (*P or p*) and *Autonomy* (*a or n*) variables.

In the second part, the general learning models are explored. Models with 2 or 3 states are investigated. Each individual sequence is considered as the replicate of the same model. The parameters of the models are obtained using the maximum likelihood estimates calculated from the observations using the EM algorithm. These statistics are provided by the R program *RHmm* (Taramasco, 2009). Models are compared using loglikelihood, AIC and BIC criteria using equations 2 and 3. Individual optimal state sequences are then obtained using the Viterbi algorithm. A goodness-of-fit test is also presented.

Results

Overall, the apprentices were autonomous in 71.4% (n = 1 375, N=1926) of the observed episodes. In 42% of the cases (n = 820), errors and/or forgetting were recorded. In 143 cases (7.4%), appropriate explanations about the activities were not provided by the apprentices. In 219 activities (11.37 %), the follow-up provided to the patient was considered incorrect. From these raw data, each of the 1,926 episodes has been classified as one instance of the observation set Ω = {*cp_n*, *cp_a*, *Cp_n*, *Cp_a*, *cP_n*, *cP_a*, *CP_n* *CP_a*}. Figure 3 presents these raw data[2] for each observation day. Learning is observable from an increase in *CP* episodes and a decrease in both *Cp_n*

---

[2] Note that only one instance of *cP_a* and none of *cP_n* have been observed. To eliminate many zero parameters in the models, the *cP_a* has been recoded as a *CP_a*.

Table 1. Number of free parameters (n_P), loglikelihood (*Log L*), *AIC* and *BIC* coefficients for models with one, two and three states.

States	n_p	Log L	AIC	BIC
1	6	-2384	4780	4813
2	13	-2250	4526	4598
3	23	-2229	4504	4631

and *cp* episodes.

Identification of the skill development model. The aim of these analyses is to determine whether or not a general skill development model emerges from this learning context. Table 1 presents the loglikelikood, AIC and BIC indexes for models from 1 to 3 states. The 3-state model has the lowest AIC but the 2-state model has the lowest BIC. As said earlier, AIC favors the selection of more complex model as compared to BIC. In such a case, the selection of a 3-state over a 2-state model might be based on theoretical or pragmatic considerations. The next sections will develop the 3-state model and will highlight that this 3-state model is useful when a researcher is interested in having more details about the transition from the unlearned to the learned state. Otherwise, the 2-state model should be preferred.

The parameters of the 3-state model are presented in Table 2. Overall, state transitions highlight a step model. The path goes from error state N_1 to intermediary N_2 state and from N_2 to N_3, the learned state. Figure 4 illustrates this transition process. The learning rate is initially slow. Transition rate from the unlearned state to the intermediary state is 2%. Once in the intermediary state, apprentices generally transit rather rapidly (30%) to the learned state. Once in the learned state, they may however regress into state 2 about 16% of the times. This suggests that the apprentices encounter new and still very difficult situations even after reaching the learned state.

For illustrative purposes, some individual predictions are shown in Figure 5. These individual predictions are obtained by applying the Viterbi algorithm to a given observed sequence for a specified state model. Transition paths for three apprentices (3, 8, and 13) are detailed. Visual inspection of apprentice 3's path shows that she remains in state 1 for some period and then rapidly transits to state 3. She very briefly visited intermediary state 2. She passes her apprenticeship. Seven other apprentices present similar learning curves. Apprentice 8 also remains in state 1 for a long period and transits to state 3. However, she regularly regresses to state 2. Therefore, at this time, her skill is not completely developed and she still needs some supervision. Three other apprentices are in a similar situation. Finally, apprentice 13 remains in the unlearned state for the whole observation period. She clearly failed her apprenticeship.

Table 3 is the observation function O and shows the nature of the hidden states. It relates each state to the observed trigrams. As expected, the learned state N_3 is made of CP_a episodes at 89%. The intermediary N_2 state is a mixture of Cp_a, CP_a and Cp_n. The sum of Cp_a (0.41) and CP_a (0.31) is 72% and suggests that the apprentices are autonomous most of the time. However, their procedural schemes still need to be improved, as 68% of the episodes are Cp_a (0.41) and Cp_n (0.27). Finally, the unlearned state N_1 is made of observations where the apprentices are not autonomous at 49%. Moreover, this state is made of cp_n, observations at 11%, suggesting that they are disoriented both from conceptual and instrumental points of views by the situations.

Next, the goodness-of-fit test is presented. It determines whether or not the model adequately predicts the data pattern as it evolves over time. Two time categories (c) will be considered here. The first is the day time category. This time category provides a preliminary overview of the fit of the model. The second is the trial time category, which will be used as the goodness-of-fit test.

Table 2. Three state step model. Transition probabilities and initial distribution (π).

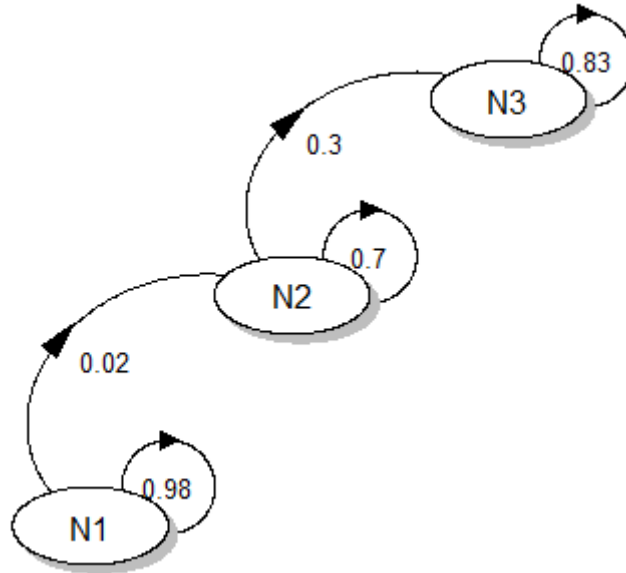| Current state	Transition to state			
	N_1	N_2	N_3	π
N_1	0.98	0.02	0.00	1.00
N_2	0.0	0.70	0.30	0.00
N_3	0.0	0.16	0.83	0.00

Figure 4. General skill development model. The hidden states N are represented by circles and transition probabilities by arrows. A 3-steps model emerges.

For the overview, predictions for the main types of observations were aggregated for each day of observation (see Figure 3). It is worth mentioning that the visual fit is rather good. The model effectively describes the evolution of observations on all days. However, HMMs can make much more precise predictions. To illustrate, a goodness-of-fit test for the CP_a observations at each observation time was performed. The hidden states at each observation time (from
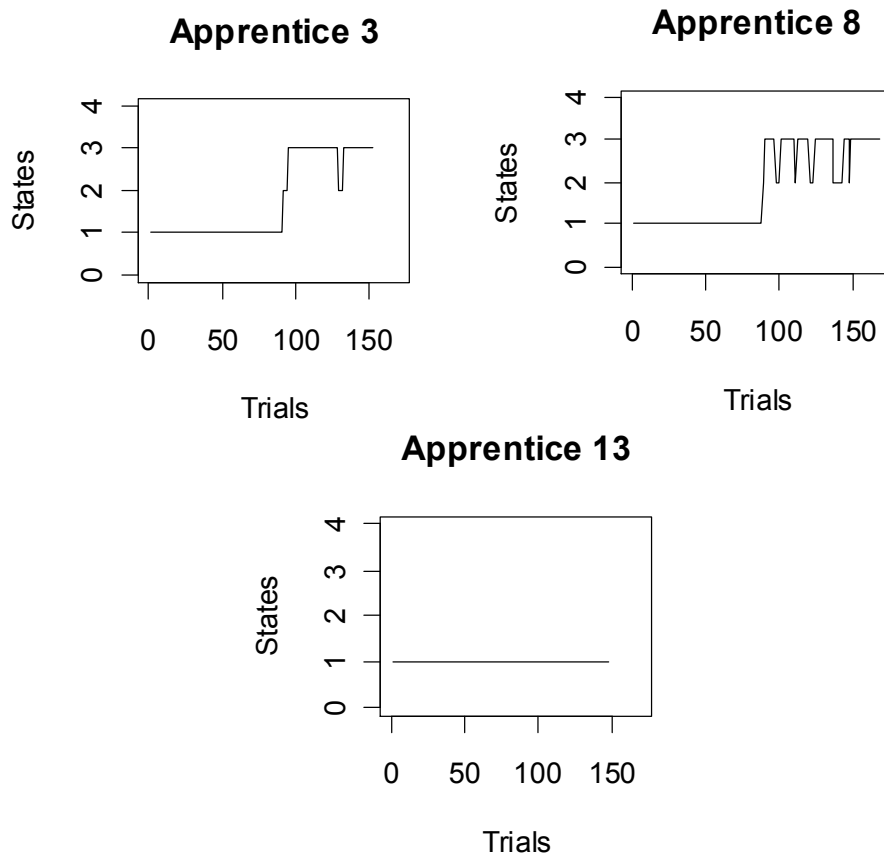

Figure 5. Learning paths for three apprentices.

Table 3. Three state step model. Conditional probability $p(O/N)$ of an observation as a function of hidden states N (skill level).

Levels	Observations				
	CP_a	Cp_a	Cp_n	cp_n	*Interpretation*
N_3	**0.89**				Is autonomous. Correctly conceptualize and execute the set of interventions.
N_2	0.31	**0.41**	.27		Is autonomous most of the times. Correctly conceptualize the sets of interventions. The interventions need some instrumental refinement.
N_1	0.35	0.16	0.34	**0.11**	Half of the times is not autonomous. Conceptualisation is adequate but interventions have important instrumental pitfalls. Constant support is needed.

For clarity, coefficients lower than 0.10 are not represented.

1 to 169) were used to compute the expected frequencies of CP_a observations. This paper presents only the test for these observations for two reasons. First, this type is closely associated with the evolution of the learned state. Second, it has a sufficient number of data available at each time period to run the χ^2 test. Figure 6 presents the results. With 169 time categories (c) and 23 free parameters (p) in the model, the χ^2 equals 72.2 with 145 df's. It fails to be significant (p < 1) and suggests that the model is adequate. There is no significant difference between the observed and predicted frequencies of CP_a episodes. Together, both the overall (visual) and the goodness-of-fit tests suggest that the model adequately describes the raw data.

Discussion

This paper proposes that Hidden Markov Models can adequately grasp the competence development process of nursery apprentices in a clinical context. The model highlights three states and a progressive step path. Autonomy and the presence of adequate conceptual and
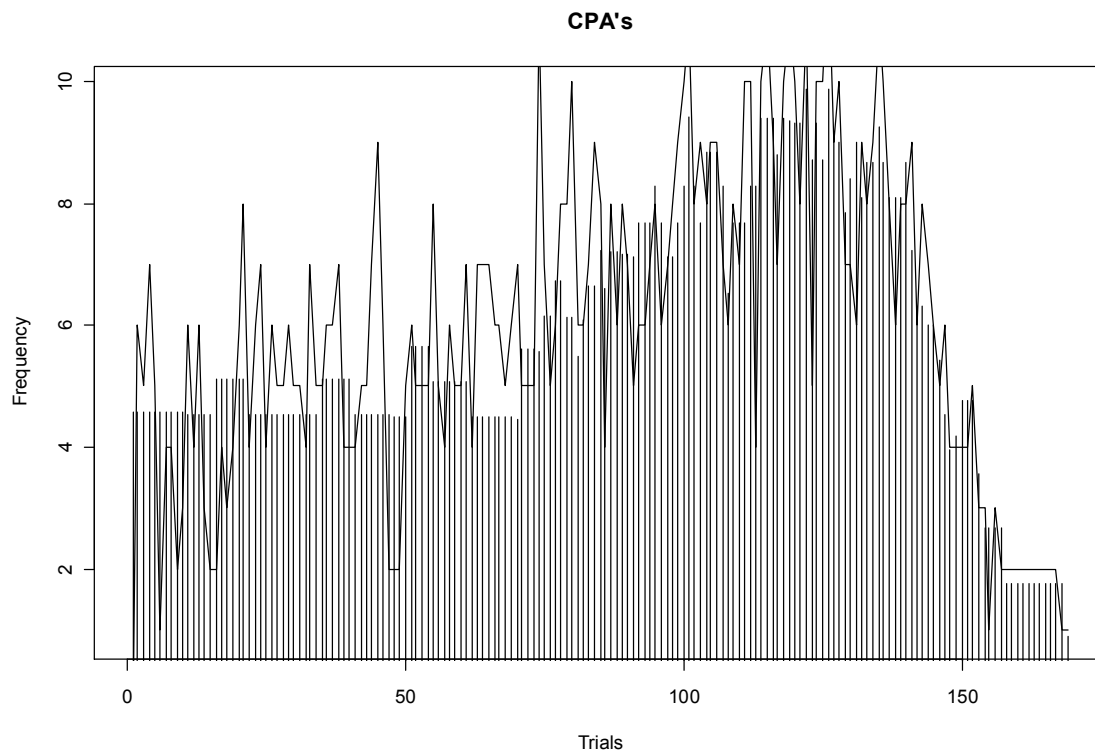

Figure 6. Observed and expected distributions of CP_a at each trial (from 1 to 169). Chisq = 72.2, df = 138 (169-30-1), p < 1. Expected are vertical bars and the continuous line, the observed

procedural schemes composed the learned state. The absence of conceptual and procedural knowledge as well as autonomy composed the unlearned state. An intermediary state with partial knowledge has also been inferred.

This 3-state model is different from the 2-state model based on success and errors as considered in implicit learning (Visser et al., 2007). The main difference is that it accounts for both conceptual and procedural aspects that are often neglected in the literature but are nonetheless important in the health care context (Lauder et al., 1999). Therefore, as new situations are encountered, the competence develops while conceptual and instrumental schemes transit in a continual series of re-representations (Anderson, 2005). The rates of progression in this step model range between 2% and 30% and fluctuate from path to path. In implicit conceptual learning, Visser et al. (2002) report comparable rates in the range of 3% to 23%.

However, there are some discrepancies in this general model that indicate that the apprentices do experience some difficulties in adapting to some unpredictable situations. First, the skill acquisition process is marked by regressions to the intermediary state. Indeed, in the learned state, there is a 16% probability of regression to the antecedent state. Moreover, from an ideal perspective, observing a learned state composed of only 89% of successes at the end of this health care curriculum might be considered problematic. It suggests that 11% of the interventions have some conceptual or instrumental pitfalls. These results are interpreted in terms of lack of transfer (Lauder et al., 1999) and needs for further supervision (Wholley & Jarvis, 2007), at least for some of the apprentices. In this respect, the model precisely portrays the progression path of each apprentice.

Conclusion

This paper suggests that a Hidden Markov Model is an adequate description of competence growth in a professional context. It extends previous work based on the analysis of success and error data (Visser et al., 2002; 2007; Wickens, 1982) and discusses some important topics, such as model specifications, parameter estimation, model selection, the Viterbi algorithm and goodness-of-fit.

In education, these models are not well known. Competences are usually inferred from series of standard items (Leighton & Gierl, 2007). It is then possible to distinguish the item difficulty from the student's ability level. However, in professional situations, cases emerge from practice and are not standard. In Markov models, competence is inferred from observations of some key variables. These models are best for describing competence changes over time. However, some caution is needed when interpreting fluctuations in these parameters, as the complexity of the situations and the apprentice ability level are confused. Hopefully, future models can disentangle these issues by recording an index of task difficulty. In summary, HMMs are flexible and provide new and insightful models for the dynamic assessment of competences in classroom and apprenticeship contexts.

References

Anderson, J. R. (2005). Human symbol manipulation within an integrated cognitive architecture. *Cognitive Science*, *29*, 313-341.

Böckenholt, U. (1999). Measuring changes: Mixed Markov models for ordinal panel data. *British Journal of Mathematical and Statistical Psychology*, *52*, 125-136.

Bruner, J. S., Goodnow, J., and Austin, G. A. (1977). A study of thinking. N.Y., Huntington: Robert E. Krieger Pub.

Duys, D. K. & Headrick, T. C. (2004). Using Markov chain analyses in counsellor education research. *Counselor Education & Supervision*, *44*, 108-120.

Eid, M., & Langeheine, R. (2007). Detecting population heterogeneity in stability and change in subjective well-being by mixture distribution models. In Anthony Ong and Manfred H. M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology*. (pp. 608-632). Oxford: Oxford University Press.

Fu, W.-T., & Anderson, J. R. (2006). From recurrent choice to skill learning: A reinforcement-learning model. *Journal of Experimental Psychology*, *135*(2), 184-206.

Harvey, L. (2009). L'échafaudage lors de la supervision en milieu professionnel : Étude des modalités et un modèle. *Mesure et évaluation en éducation*, *32*(1), 55-83.

Harvey, L., & Anderson, J. R. (1996). Transfer of declarative knowledge in complex information processing domains. *Human-Computer Interaction*, *11*, 69-96.

Harvey, L., & Barras, D. (2008). Transfert des compétences et construction d'un langage d'action en situation de compagnonnage professionnel. [Transfer of skills and action language acquisition in professional apprenticeship]. *Revue des sciences de l'éducation, 34*, 665-687.

Hélie, S. (2006). An introduction to model selection : Tools and algorithms. *Tutorials in Quantitative Methods for Psychology*, *2*(1), 1-10.

Langeheine, R., Stern, E., & van de Pol, F. (1994). State mastery learning: Dynamic models for longitudinal data. *Applied Psychological Measurement*, *18*(3), 277-291.

Lauder, W, Reynolds, W., & Angus, N. (1999). Transfer of knowledge and skills: some implications for nursing and nurse education. *Nurse Education Today*, *19*, 480-487.

Leighton, J, O., & Gierl, M, J. (2007). *Cognitive Diagnostic Assessment for Education: Theory and Applications*

Retrieved 18 March 2009, from http://www.myilibrary.com/Browse/open.asp?ID=90996&loc=Cover

Littman, M. L. (2010). A tutorial on partially observable Markov decision processes. *Journal of Mathematical Psychology*, *53*(3), 119-125.

Miller, G. A. (1952). Finite Markov processes in psychology. *Psychometrika*, *17*(2), 149-167.

Ministère de l'éducation, des loisirs et du sport (2004). *180.A0 Soins infirmiers* [En ligne]. Disponible le 13 janvier 2008: http://www.meq.gouv.qc.ca/ens%2Dsup/ens%2Dcoll/cahiers/program/180a0.asp

McCoach, D. B., & Black, A. C. (2008). Evaluation of model fit and adequacy. In Ann. A. O'Connell and D. Betsy McCoach (eds.), *Multilevel modeling of educational data.* (pp. 245-273). *Greenwich* : Information age Publishing Inc.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139-154

Pata, K., Lehtinen, E., & Sarapuu, T. (2006). Inter-relations of tutor's and peers' scaffolding and decision making discourse acts. *Instructional Science*, *34*, 313-341.

Potgieter, M., Harding, A., & Engelbrecht, J. (2008). Transfer of algebraic and graphical thinking between mathematics and chemistry. *Journal of Research in Science Teaching*, *45*, 197-208.

Taramasco, O. (2009). Hidden Markov Models simulations and estimations (version 1.3.1). http://www.r-project.org, http://r-forge.r-project.org/projects/rhmm/

Sabah, G. (2002). Schéma, Schème. [Schemes, Schemas]. In G. Tiberghien (Ed.): *Dictionnaire des sciences cognitives.* Paris: Armand Colin.

Singley, M. K., et Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Steiner, T. E., & Greeno, J. G. (1969). An analysis of some conditions for representing N state Markov processes as general all or none models. *Psychometrika*, *34*, 461-487.

Visser, I., Maartje E. J. Raijmakers, Han L. J., & van der Maas, H. L. J. (2009). Hidden Markov Models for Individual Time Series. In Jaan Valsiner, P.C.M. Molenaar, Maria C.D.P Lyra, Nandita Chaudhary (eds.), *Dynamic Process Methodology in the Social and Developmental Sciences.* (269-289). N. Y : Springer-Verlag.

Visser, I, Schmittmann, V. D., & Raijmakers, M. E. J.. (2007). Markov process models for discrimination learning. In K. van Montfort, H. Oud & A. Satorra (Eds.). *Longitudinal models in the behavioral and related sciences*.(pp.337-365). Mahwah: Lawrence Erlbaum Associates.

Visser, I., Raijmakers, M. E. J., & Molenaar, P. C. M. (2002). Fitting hidden Markov models to psychological data. *Scientific Programming*, *10*, 185-199.

Vosniadou, S., (2007)**.** The cognitive-situative divide and the problem of conceptual change, *Educational Psychologist*, *42*(1), 55-66.

Wholley, N. N., & Jarvis, Y. (2007). Situated cognition and cognitive apprenticeship: A model for teaching and learning clinical skills in a technology rich and authentic learning environment. *Nurse Education Today*, *2, 73-79.*

Wickens, D. W. (1982). *Models for behavior: Stochastic processes in psychology*. San Francisco: W. H Freeman and Company.

Weingart, L. R., Prietula, M. J., Hyder, E. B., & Genovese, C. (1999). Knowledge and the sequential processes of negotiation: A Markov chain analysis of response-in-kind. *Journal of Experimental Social Psychology*, *35*, 366-393.