HELSINKI INSTITUTE FOR INFORMATION TECHNOLOGY

EDUFORM – A TOOL FOR CREATING ADAPTIVE QUESTIONNAIRES

Miikka Miettinen, Petri Nokelainen, Jaakko Kurhila,
Tomi Silander and Henry Tirri

January 14, 2004

HIIT
TECHNICAL
REPORT
2004–7

EDUFORM – A TOOL FOR CREATING ADAPTIVE QUESTIONNAIRES

Miikka Miettinen, Petri Nokelainen, Jaakko Kurhila, Tomi Silander and Henry Tirri

Helsinki Institute for Information Technology HIIT
Tammasaarenkatu 3, Helsinki, Finland
PO BOX 9800
FI-02015 TKK, Finland
`http://www.hiit.fi`

HIIT Technical Reports 2004–7
ISSN 1458-9478
URL: http://cosco.hiit.fi/Articles/hiit-2004-7.pdf

Copyright © 2004 held by the authors

NB. The HIIT Technical Reports series is intended for rapid dissemination of results produced by the HIIT researchers. Therefore, some of the results may also be later published as scientific articles elsewhere.

EDUFORM – A Tool for Creating Adaptive Questionnaires

Miikka Miettinen[1], Petri Nokelainen[1], Jaakko Kurhila[2], Tomi Silander[1], Henry Tirri[1]

firstname.lastname@helsinki.fi

1) Helsinki Institute for Information Technology
Complex Systems Computation Group
P.O. Box 9800, 02015 HUT, Finland
Tel: + 358 9 850 11576, Fax: + 358 9 694 9768

2) University of Helsinki
Department of Computer Science
P.O. Box 26, 00014 Univ. of Helsinki, Finland
Tel: + 358 9 1911, Fax: + 358 9 191 44441

Abstract: Questionnaire data has many important uses, but is laborious for the subjects to provide. EDUFORM tries to alleviate this problem by enabling the creation of adaptive on-line questionnaires. The idea is to build a probabilistic model from previously gathered data, and employ it for predicting the profiles of new users on the basis of a subset of the questions in the original questionnaire. The questions presented to each individual are selected adaptively in order to minimize the number of answers needed. Empirical evaluations suggest that 85-90% profiling accuracy can be achieved, while the number of answers is reduced by 30-50%.

Introduction

The information needs involved in organizing effective education are significant. Accurate knowledge of the students' interests, preferences, and motivation is important both for the daily activities of educational institutions and for longer-term research and development efforts. In addition, computer technology enables such information to be used for the immediate benefit of the students. Self-assessment tools can be developed to provide analyses of learning styles or metacognitive skills, and adaptive systems to adjust the content or presentation of the material to individual needs. The problem is that nearly all of the interesting and useful information has to be provided explicitly by the students, which easily leads to excessive use of questionnaires. Besides being undesirable in itself, the tedious and sometimes frustrating answering process associated with long questionnaires is likely to reduce the reliability of the acquired data.

In order to address this problem, we have developed EDUFORM, a generic tool for creating adaptive multiple-choice questionnaires. The idea behind EDUFORM is to build a model from previously gathered data and employ it for profiling new users on the basis of a subset of the questions in the original questionnaire. Furthermore, the questions and the order in which they are presented are chosen adaptively on the basis of the previous answers of the particular individual. Our empirical evaluations suggest that good profiling accuracy can be achieved with a significantly reduced number of answers.

Modeling approach

The typical way of using questionnaires is to look for characteristic answer profiles, and utilize them as abstractions in subsequent theoretical or practical analyses. As a result, the most relevant form of adaptation involves making predictions about the user's profile. Although the predictive model can in principle be derived in a theory-driven manner and coded manually, we have adopted a data-driven viewpoint, which means that the model is constructed from data gathered previously with the same questionnaire. This leads to the distinction of two phases in the use of EDUFORM: the *Profile creation phase*, where the set of characteristic profiles is captured in a model, and the *Query phase*, where the model is used for adaptive questioning. The design is generic and allows the application of any type of predictive model suitable for the task. We have adopted the Bayesian approach (Bernardo and Smith, 2000) and use the language of probability distributions to describe the profiles.

In practice, the profiles are constructed by dividing the data vectors to a number of mutually exclusive groups and summarizing the contents of each group in a probability distribution. The details of the procedure are described in (Kontkanen et al. 1996) and (Tirri et al. 1996), but the underlying intuitive idea can be illustrated briefly as follows. Each profile is a prototype, which can be employed for creating a more compact representation of the data. In order to describe an individual data vector, it is sufficient to specify the closest prototype and list the differences between the expected and observed values. Alternative choices of the prototypes can be evaluated on the basis of the amount of information needed to describe the entire data set: the more representative the prototypes are, the fewer differences need to be listed one-by-one.

The resulting model serves two different purposes. On the one hand, it is a useful representation of statistical regularities in the data. The answer distributions associated with each profile can be extracted from the model, analyzed, and compared to each other. Since the profiles are based only on the data and some general assumptions of the model class, they also constitute an empirical test for the theory or hypotheses that guided the design of the questionnaire. On the other hand, the model is suitable for the kind of prediction needed in the *Query phase*, as will be explained later in this paper.

Figure 1: The user interface of EDUFORM.

EDUFORM

Even though EDUFORM is an electronic questionnaire on-line, it resembles traditional questionnaires on paper (see Figure 1). The questions appear inside a fixed size rectangular area with a navigation bar at the bottom. Only 3-5 questions are shown simultaneously to allow the order of the remaining questions to be adapted dynamically and to eliminate the need for scrolling. The arrows on the right side of the navigation bar allow the user to move to the next or to the previous set of questions. An answer can be supplemented with a free-form comment by clicking the pencil icon beside the radio buttons. Once a comment has been written, the pencil changes into a paper, as in the middle of Figure 1. Clicking the button marked with the pie chart icon shows the user's current profile. When the profile is known with sufficient certainty, the user can skip the remaining questions by clicking the button with the cross on it. On the left side of the navigation bar is a progress indicator showing an estimate of the proportion of questions left. When the mouse pointer moves on top of a button or the progress indicator, the name of the button or the current value of the indicator is shown as a tooltip. Because of the simplicity of the interface, there is no need for a separate help screen.

Adaptation in EDUFORM

In the *Query phase*, we want to find out the profile of the user as efficiently as possible. The profile is represented by a probability distribution for the groups identified in the *Profile creation phase*. As the user answers the questions, some of the groups become much more likely than others, and one of them often reaches almost 100% probability rather quickly. EDUFORM takes advantage of this characteristic pattern by optimizing the order in which the questions are presented, and offering the user a chance to quit once sufficient certainty about the profile has been achieved.

At any point in time, the most informative set of questions to ask next is the one that is expected to change the profile distribution most. EDUFORM searches for this set by maximizing the *Kullback-Leibler distance* (Cover and Thomas 1991) between the current distribution and the distribution that would be expected if answers to a particular set of additional questions were received. The first questions are the same for everybody, but after that the selection depends on the previous answers of each individual. Therefore, adaptation in EDUFORM is based on continuous assessment of the expected information gain, rather than being limited to a small number of hard-coded paths.

The purpose of this technique is to minimize the number of answers needed to find out the user's profile. Additional questions can be omitted entirely once a sufficient degree of certainty has been achieved. In the current experimental version of EDUFORM, the termination criterion is defined by setting a limit, which the most probable group in the profile has to exceed. A value within 75-85% seems to be suitable in most cases. It is also possible to specify an additional requirement regarding the stability of the profile. For example, it may be stated that the most probable group has to stay above the limit for two successive sets of answers.

Box 1: Prediction and adaptation with a Bayesian finite mixture model.

If Q denotes the filled-in questionnaire, each group G_i identified in the data can be described as a mechanism that assigns a probability $P(Q \mid G_i)$ to the questionnaire. The set of groups $G = (G_1, G_2, \ldots, G_K)$, together with their relative sizes $s = (s_1, s_2, \ldots, s_K)$, define a finite mixture (Titterington et al. 1985) that can be treated as a probability model

$$P(Q \mid G, s) = s_1 P(Q \mid G_1) + s_2 P(Q \mid G_2) + \ldots + s_K P(Q \mid G_K).$$

As the adaptive questionnaire is being completed, the probabilities in the model are updated to reflect the new information gained from the answers. The model allows us to calculate the probabilities of the possible answers to the unanswered questions (Q_U) on the basis of the answered questions (Q_A):

$$P(Q_U \mid Q_A, G, s) \propto P(Q_U, Q_A \mid G, s) = P(Q \mid G, s).$$

We can also keep track of the probability of each particular group G_i in the profile distribution. If we denote by g the event that the user belongs to the group G_g,

$$P(g \mid Q_A, G, s) \propto P(g, Q_A \mid G, s) = P(g \mid G, s)\, P(Q_A \mid g, G, s) = s_g\, P(Q_A \mid G_g).$$

The most informative subset of the unanswered questions is determined by

$$\operatorname*{argmax}_{Q_X} \sum_{q_X \in \mathrm{ans}(Q_X)} P(q_x \mid Q_A, G, s) \sum_{i=1}^{K} P(g_i \mid Q_A, G, s) \log \frac{P(g_i \mid Q_A, G, s)}{P(g_i \mid Q_A, q_x, G, s)},$$

where Q_X is the subset being considered and $\mathrm{ans}(Q_X)$ is the set of its possible answer combinations. The inner sum is the Kullback-Leibler distance between the current profile distribution and the distribution that would result if the user gave the answers q_x. The outer sum adds up the contributions of the individual answer combinations, weighing them by their probabilities.

The mathematics underlying the adaptation mechanism is summarized in Box 1. It should be noted that several adjustments could be tried to improve the results, and our approach is not the only way of making adaptive questionnaires by means of statistical learning. For example, Johnson and Albert (1999, 191) have proposed an alternative technique based on the estimation of item specific model parameters.

Figure 2 shows the format in which the data is saved. The first column identifies the person. In this particular case, a unique identification string has been created from the questionnaire name ("demo") and a counter. The questions appear in the same order as they were presented to the user. Question numbers are in the second column. The remaining columns contain the probabilities of the possible answers. If the user has actually answered the question, one of the probabilities is 1 and the rest are 0. Probability distributions for the omitted questions are calculated from the model and saved in the same file. In Figure 2, the first four questions have been answered by the user, and the last two rows are predictions. Additional data includes comments, the final profile, and a log of mouse clicks. The main purpose of the log is to record the time used for answering various parts of the questionnaire, but it may also be helpful for identifying ambiguous questions or making detailed analyses of differences between groups of users.

```
demo-1  33    0.0    1.0    0.0    0.0    0.0
demo-1  15    0.0    1.0    0.0    0.0    0.0
demo-1  10    0.0    0.0    0.0    1.0    0.0
demo-1  27    0.0    0.0    0.0    1.0    0.0
demo-1  5     0.0149 0.0292 0.1225 0.2392 0.5939
demo-1  11    0.0084 0.0086 0.0422 0.2451 0.6954
```

Figure 2: Format of the saved data.

Empirical results

Perhaps the most important question to ask when judging the value of EDUFORM is whether or not it actually works. The number of answers needed for reliable profiling should be significantly smaller than the total number of questions in the questionnaire. We would also like the users to take advantage of the adaptivity and quit when they are offered a chance to do so.

In order to evaluate the predictive performance of EDUFORM, we simulated the operation of the adaptive questionnaire using complete data. The models were constructed from 200 randomly selected cases in each data set, and the remaining test cases were supplied to the models exactly as they would have been received during the course of adaptive questioning. The number of answers given before the fulfillment of the termination criteria was recorded, and the group predicted at that point was compared to the group assigned after the remaining answers had been supplied to the model. If the predicted group did not match the final group, the prediction was recorded as an error.

Table 1 shows the main results of the simulation. Two different data sets were available from a questionnaire (Ruohotie 2001) with four sections: "Learning and motivation" (Motiv in Table 1), "Study habits" (Habits), "The quality of teaching" (Teaching), and "The effects and outcomes of education" (Effects). Although the sections measure complementary aspects of the same educational setting, they are in the present context best thought of as separate questionnaires. The last data set (Motprof) is from a questionnaire designed for identifying motivational profiles. The second and third columns contain the number of groups identified during model construction and the total number of questions in the questionnaire. The average proportion of questions needed for predicting the group of a test case is in the column labelled "Questions asked". The next two columns contain the standard deviation of the number questions asked and the proportion of test cases for which the final group differed from the group predicted upon the fulfilment of the termination criteria.

Data set	Groups	Number of questions	Questions asked	Standard dev. of quest. asked	Errors	Number of test cases
Motiv 1	4	28	62%	22%	10%	260
Motiv 2	4	28	65%	22%	15%	357
Habits 1	5	40	62%	22%	15%	260
Habits 2	5	40	48%	21%	13%	357
Teaching 1	5	23	67%	21%	13%	260
Teaching 2	5	23	53%	24%	15%	357
Effects 1	5	25	61%	22%	14%	260
Effects 2	5	25	45%	23%	14%	357
Motprof	6	34	70%	21%	15%	498

Table 1: Predictive performance of EDUFORM.

As can be seen in Table 1, on average 50-70% of the questions had to be asked to achieve an error rate of 10-15%. Every data set contained a few exceptional cases for which 100% or only 15-30% of the answers were needed, but the standard deviations were consistently within 20-25% of the total number of questions in the questionnaire.

The trade-off between the number of answers and the number of errors can be altered by adjusting the termination criteria. The more uncertainty we are willing accept in the profile, the fewer questions need to be asked. Figure 3 shows the effect of additional answers in the Motprof data set. On the horizontal axis we have the number of answers given, and on the vertical axis the average Kullback-Leibler distance between the predicted profile and the final profile. By setting the termination criteria to appropriate values, questioning can be stopped approximately at the desired point along the line.

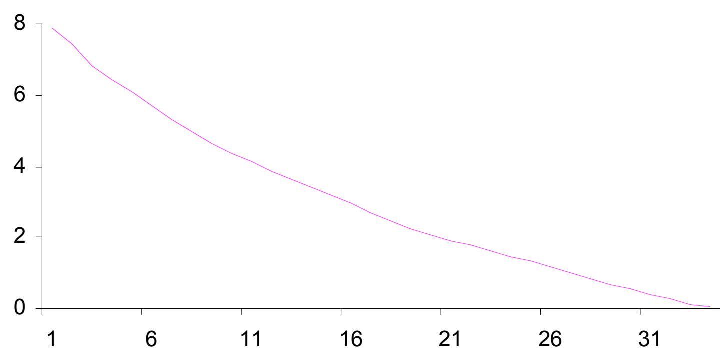

Figure 3: Reduction in the distance between the predicted and the final profile.

At the time of writing, two data sets have been gathered with the adaptation mechanism turned on. The same questionnaires were used as in the simulation study described above. Of particular interest for the present purpose is the attitude of the users towards prediction. When their predicted profile satisfied the termination criteria, they were asked if they want to quit or refine the profile by answering the remaining questions. They could also quit after answering only some of the additional questions. The decision to quit or continue can be seen as a reflection of the user's opinion about the usefulness of the adaptivity. Those who took advantage of the possibility of skipping questions probably considered it a helpful feature, whereas the others either did not mind answering all questions or had doubts about the reliability of the predictions.

The results are summarized in Table 2. The first four questionnaires were parts of the same study, and were completed sequentially during one session. The subjects were students from a teacher training programme in the Finnish Polytechnic Institute. In the other study ("Motprof"), motivational characteristics of engineering students from the Helsinki University of Technology were examined. The second column contains the proportion of users who quitted before answering all questions. Unfortunately, it seems that the adaptivity was not appreciated as much as we thought it would be. The third column shows the number of questions answered by the students who did quit before the end. The second part of the first study ("Habits") was the longest one with 40 questions. The proportion of the answered questions is high because many students gave a few more answers after they had the first chance to quit, but got tired before reaching the end. Taking this into account, the predictive performance of the models was at the same level with the simulation results.

Questionnaire	Allowed prediction	Questions answered	Total number of cases
Motiv	11%	64%	66
Habits	35%	82%	66
Teaching	20%	61%	66
Effects	17%	68%	66
Motprof	26%	61%	478

Table 2: The adaptivity of EDUFORM in real use.

Conclusions

EDUFORM is a tool for increasing the efficiency of questionnaires with adaptation and prediction. The underlying software is independent of questionnaire content. This domain independence opens up the possibility of using EDUFORM for more than just a single purpose. For example, EDUFORM questionnaires could be applied to assessing individual differences on-line in order to provide support for studies in a virtual or traditional university. A questionnaire giving personalized tips for more effective studying would be appropriate support material for student self-evaluation.

An adaptive questionnaire created with EDUFORM could also be used as a test for students. Testing the students' knowledge with adaptive questioning is not a novel idea. However, in the standard approach the system adapts directly to the knowledge of the student. When using an EDUFORM questionnaire as a test, adaptation means the optimization of the length of the test. In other words, the goal is to provide the teacher or evaluator with sufficient information about the students' progress asking as few questions as possible.

Because of the particular approach to modeling and adaptation, EDUFORM could also be used as a tool for creating user profiles for adaptive educational systems. Sufficient knowledge of the characteristics of the user is a necessary prerequisite for effective adaptation. Some systems are able to accumulate useful data during the course of their interaction with the user, but additional input must almost always be provided explicitly (Brusilovsky 2001). EDUFORM could be employed for gathering this information efficiently or creating probabilistic user profiles for direct application in the other system.

The current version of EDUFORM is suitable for testing and experimentation, but it is not accessible to a non-technical user. The creation of new adaptive questionnaires is based on a command line interface and manual editing of configuration files. In addition, the modeling component requires the previously gathered data to be supplied in a particular format, and does not currently include any conversion utilities. There is nothing in principle, however, which would prevent the development of a self-contained and user-friendly software package. Most of the required features would be relatively simple supplements providing more convenient access to the existing core functionality. Whether or not there would be sufficient demand for such a package is still an open issue, which depends primarily on the applicability of EDUFORM to real-world use.

Acknowledgements

This work was supported in part by the Academy of Finland under the Prima and Prose projects.

References

Bernardo, J. & Smith, A. (2000). *Bayesian Theory*. John Wiley & Sons: New York. 2nd ed.

Brusilovsky, P. (2001). Adaptive hypermedia. *User Modeling and User Adapted Interaction*, Ten Year Anniversary Issue (Alfred Kobsa, ed.) 11 (1/2), pages 87-110.

Cover, T. & Thomas J. (1991). *Elements of Information Theory*, John Wiley & Sons, New York, NY.

Johnson, V. & Albert, J. (1999). *Ordinal Data Modeling*. Springer: New York.

Kontkanen, P., Myllymäki, P. & Tirri, H. (1996). Predictive Data Mining with Finite Mixtures. In *Proceedings of The Second International Conference on Knowledge Discovery and Data Mining*, pages 176-182. Portland, OR, August 1996.

Ruohotie, P. (2002). Motivation and Self-regulation in Learning. In H. Niemi & P. Ruohotie (Eds.), *Theoretical Understandings for Learning in the Virtual University*, pages 37-72. Research Centre for Vocational Education, University of Tampere.

Tirri, H., Kontkanen, P. & Myllymäki, P. (1996). Probabilistic Instance-Based Learning. In L. Saitta, *Machine Learning: Proceedings of the Thirteenth International Conference*, pages 507-515. Morgan Kaufmann Publishers: San Francisco.

Titterington, D. and Smith, A. & Makov, U. (1985). *Statistical Analysis of Finite Mixture Distributions*. John Wiley & Sons: New York.