Developing a Bayes-net based student model for an External Representation Selection Tutor

Beate Grawemeyer and Richard Cox
Representation & Cognition Group
Department of Informatics, University of Sussex, Falmer, Brighton BN1 9QH, UK

Abstract. This paper describes the process by which we are constructing an intelligent tutoring system (ERST) designed to improve learners' external representation (ER) selection accuracy on a range of database query tasks. This paper describes how ERST's student model is being constructed - it is a Bayesian network seeded with data from experimental studies. The studies examined the effects of students' background knowledge-of-external representations (KER) upon performance and their preferences for particular information display forms across a range of database query types.

Keywords. Student modeling, External representations, Bayesian networks

1. Introduction

Successful use of external representations (ERs) depends upon the skillful matching of a particular representation with the demands of the task. Good ER selection requires, *inter alia*, knowledge of a range of ERs in terms of a) their semantic properties (e.g. *expressiveness*), b) their functional roles (e.g. [4],[1]) together with information about the 'applicability conditions' under which a representation is suitable for use [7].

Our aim is to build ERST - an ER selection tutor. We conducted a series of empirical studies (e.g. [6]), that have provided data for ERST's student model and it's adaptation mechanism. This paper extends the work by investigating the effect of learners' background knowledge of ERs (KER) upon information display selection across a range of tasks that differ in their representation-specificity. In the experiments, a prototype automatic information visualization engine (AIVE) was used to present a series of questions about information in a database. Participants were asked to make judgments and comparisons between cars and car features. Each participant responded to 30 questions, of which there were 6 types, e.g. identify; correlate; quantifier-set; locate; cluster; compare negative. Participants were informed that to help them answer the questions, the system would supply the needed data from the database. AIVE then offered participants a choice of representations of the data. They could choose between various types of ERs, e.g. set diagram, scatter plot, bar chart, sector graph, pie chart and table. The ER options were presented as an array of buttons each with an icon depicting, in stylized form, an ER type (bar chart, scatter plot, pie chart, etc). When the participant made his or her choice, AIVE then instantiated the chosen representational form with the data needed to answer the task and displayed a well-formed, full-screen ER from which the participant could read-off the information needed to answer the question. Having read-off the information, subjects indicated their response via on-screen button selections (i.e. selecting one option out of a set of possible options). Note that each of the 30 questions could (potentially) be answered with any of the ER display types offered. However, each question type had an 'optimal' ER. Following a completed response, the participant was presented with the next question in the series of 30 and the sequence was repeated. The data recorded were: the randomized position of each representation icon from trial to trial; user's representation choices (DSA); time to read question and select representation (DSL); time to answer the question (DBQL); responses to questions (DBQA). Further details about the experimental procedure are provided in [6].

Prior to the database query tasks, participants were provided with 4 different types of KER pre-tests [5]. These tests consisted of a series of cognitive tasks designed to assess ER knowledge representation at the perceptual, semantic and output levels of the cognitive system. A large corpus of external representations (ERs) was used as stimuli. The corpus contains 112 ER examples. The decision task (ERD) was a visual recognition task requiring real/fake decisions[1]. The categorisation task (ERC) assessed semantic knowledge of ERs - subjects categorised each representation as 'graph or chart', or 'icon/logo', 'map', *etc*. In the functional knowledge task (ERF), subjects were asked '*What is this ER's function*'*?*. In the naming task (ERN), for each ER, subjects chose a name from a list. E.g.: 'venn diagram', 'timetable', 'scatterplot', 'Gantt chart', 'entity relation (ER) diagram', *etc* [5].

2. Results and Discussion

The simple bivariate correlations between KER and AIVE tasks for display selection accuracy (DSA), database query answering accuracy (DBQA), display selection latency (DSL) and database query answering latency (DBQL) were: Three of the 4 KER tasks correlated significantly and positively with DBQA (ERD $r=.46$, $p<.05$; ERC $r=.60$, $p<.01$; ERF $r=.66$, $p<.01$); Two KER tasks correlated significantly and positively with DSA (ERC $r=.57$, $p<.01$; ERF $r=.57$, $p<.01$); DBQA correlated significantly and positively with DSA ($r=.30$, $p<.01$); There is a significant negative correlation between DBQA and DBQL ($r=-.28$, $p<.01$); DSA is significantly negatively correlated with DSL ($r=.-17$, $p<.01$); There is a significant negative correlation between DSA and DBQL ($r=-.32$, $p<.01$); DSL and DBQL are significantly positively correlated ($r=.30$, $p<.01$).

The results showed that task performance on three of the KER tasks are better predictors of DBQA performance than DSA. The selection latency results show that a speedy selection of a display type in AIVE is associated with a good display-type choice. This implies that students either recognise the 'right' representation and proceed with the task or they procrastinate and hesitate because of uncertainty about which display form to choose. Less time spent responding to the database query question is associated with a good display-type choice and correct query response. This suggests that the selection and database query latencies may be used in ERST's student model as predictors of students' ER expertise.

[1]Some items in the corpus are invented or chimeric ERs.

Using the experimental data, a Bayesian network [8] was constructed for ERST's student model. Bayesian networks have been applied successfully in ITS (e.g. [2]) and are suitable for recognizing and responding to individual users, and they can adapt to temporal changes. The network will monitor and predict users' ER selection preference patterns within and across query types. It will relate query response accuracy and latencies to particular display selections and select query/display combinations to 'probe' an individual user's degree of 'graphical literacy'. The empirical data is used to instantiate values in the conditional probability tables (CPTs) at each node of the model. The network will then dynamically adjust the CPT values and evolve individualised models for each of its users as they interact with the system. The student model will drive ERST's educational interventions (by hinting or advising) or by 'hiding' inappropriate display forms. The aim is for ERST to be able to generate ER-to-task matching situations that will function as 'probes' of an individual student's knowledge. ERST will be able to interrupt if too much time is spent on selecting a representation (after learning individual's selection display selection latency patterns). It will acquire a basis for recommending the most appropriate display(s) and for varying the range of 'permitted' displays as a function of each task's ER-specificity. If a user manifests a particularly high error rate for particular task/ER combinations, then ERST will be able to offer clarification of, e.g. the functionality of that particular ER. At early stages of learner-system interaction, ERST's adaptiveness will be limited to attempts to offer only display choice options that it believes lie within the learner's 'representational repertoire'. After more extensive learner-system interactions the student model will be more established. At that point ERST may be able to make firmer recommendations to its user and may choose to directly tutor ER-to-task matching skills in the case of ERs for which the student's knowledge appears to be weak.

References

[1] P C-H. Cheng: Functional roles for the cognitive analysis of diagrams in problem solving. In G.W. Cottrell (Eds.): *Proceedings of the 18th Annual Conference of the Cognitive Science Society*, Mahweh, NJ: Lawrence Erlbaum Associates (1996), 207–212.

[2] C. Conati, A. Gertner and K. VanLehn: Using Bayesian networks to manage uncertainly in student modeling. *User Modeling & User-Adapted Interaction*, **12(4)**, (2002) 371–417.

[3] R. Cox: Representation construction, externalised cognition and individual differences. *Learning and Instruction*, **9** (1999), 343–363.

[4] R. Cox and P. Brna: Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial Intelligence in Education*, **6(2)** (1995), 239–302.

[5] R. Cox, P. Romero, B. du Boulay, and R. Lutz: A cognitive processing perspective on student programmers' 'graphicacy'. In A. Blackwell, K. Marriott & A. Shimojima (Eds.): *Diagrammatic Representation & Inference. Lecture Notes in Artificial Intelligence*, Berlin, Springer, (2004).

[6] B. Grawemeyer and R. Cox: A Bayesian approach to modelling user's information display preferences. To be published In: *UM 2005 User Modeling: The Proceeding of the Tenth International Conference*, Berlin, Springer (2005).

[7] L.R. Novick,, S.M. Hurley and M. Francis: Evidence for abstract, schematic knowledge of three spatial diagram representations. *Memory & Cognition*, **27(2)**, (1999), 288–308.

[8] J. Pearl: *Probabilistic reasoning in intelligent systems: Networks of Plausible Inference*. Morgan Kaufmann (1988).