

Analyzing Sequential Data in Computer-Supported Collaborative Learning

Dietmar Janetzko¹, Frank Fischer²

¹Institute of Computer Science and Social Research, University of Freiburg, Germany

dietmar@cognition.iig.uni-freiburg.de

²Institute of Psychology, University of Tübingen, Germany

frank.fischer@iwm-kmrc.de

ABSTRACT

Representations and changes between them play a major role in cognitive development (e.g., Vosniadou, & Brewer, 1992) and education (e.g., Hewson, Beeth, & Thorley, 1998). By definition, change of representations is also indispensable for collaborative work since a common understanding or shared knowledge can only be achieved by a partial convergence of the knowledge structures of the collaborating subjects. This article presents and discusses knowledge tracking (KT), viz., an approach to analyze cognition on the basis of symbolic sequential data. We present and discuss the methodological aspects of KT and delineate the Web-based computer program (knowledge tracking engine, KTE) set up to run KT-analyses (<http://www.knowledge-tracking.com>). An empirical study in collaborative learning is taken to exemplify the usage of KT in analysis of computer supported collaboration.

Keywords

Shared knowledge, network representation of cognition, probabilistic model, change of representation, sequential data

KNOWLEDGE TRACKING

Knowledge tracking (KT) is a psychometric method that carries out a diagnosis of cognitive representations. Knowledge tracking can be used in a *confirmative* or in a *generative* mode. The former provides a rationale to decide which of some candidate theories (relational structures) explains a sequence of data best.⁽¹⁾ The latter may be pursued to generate a relational structure on the basis of some start-up structures such that the newly generated structure fits to the data best (Janetzko, 2000).

The Data: Sequence of Concepts. The input of data required by knowledge tracking is a sequence of symbols or concepts (e.g., the sequence of the concepts CAT, DOG, FISH, MOUSE etc.) that refer to the sequence of states in a Markov process. This kind of data may be easily obtained in studies of computer supported collaboration or think aloud studies.

The Theory: Relations and Structures. Knowledge tracking needs a theory to analyze sequences of symbolic data. To specify a theory we have to select one or many simple semantic relations (e.g., *x is-a y* or *x communicates to y*). On the basis of a relation we may then add a set of concepts that are taken to instantiate the relations. We end up with relational structures. A very simple relational structure can be described in a Lisp-like notation as (is-a (MOUSE MAMMAL) (HORSE MAMMAL) (SHARK FISH) (HERRING FISH) (FISH VERTEBRATE) (MAMMAL VERTEBRATE)).

Calculating Goodness of Fit Scores. In KT, the theory, viz., one or many relational structures, is taken to calculate goodness of fit scores on the basis of sequences of symbolic data. A goodness of fit score describes how well a sequence of symbolic data can be explained by a relational structure.

The Knowledge Tracking Engine: Using Knowledge Tracking via the WWW

The web-site <http://www.knowledge-tracking.com> allows visitors to carry out remote analyses of data on the basis of the knowledge tracking engine (KTE). Users may upload data and theories and use KT either in the confirmative or in the generative mode.

USAGE OF KNOWLEDGE TRACKING TO ANALYSE SHARED KNOWLEDGE

Recent work in CSCL indicates that *knowledge convergence* in collaborative learning might be an important factor for conceptual change (Roschelle, 1995). Knowledge convergence could be defined as the increasing similarity of subjects with regard to process and outcome of knowledge construction (Fischer & Mandl, 2001). We investigated, to what extent collaborators converge with respect to process and outcomes of problem-oriented collaborative learning in text-based computer-mediated communication. Moreover, we compared two different instructional conditions with regard to their effects on knowledge convergence. Both conditions were structured by a specific type of collaboration script (e.g.,

⁽¹⁾ By *explanation* we refer to the theory-based prediction of data.

O'Donnell, 1996): A *macroscript* mainly sets a common goal and prescribes major steps. In contrast, a *microscript* is more fine-grained in regulating the learners interaction.

Participants and design. We re-analyzed data of 18 subjects from a recent study (Reiserer, Ertl, Weinberger, Fischer, Mandl, & Jahn, in press). Subjects worked together in distributed groups of three learners (triads) in one of the two different script conditions of a one-factorial design.

Procedure. Learners worked on standard personal computers with 400Mhz and 128 MB RAM located in separated places. They communicated via text-based CMC facilities. The main components of the learning environment were a text-based computer conferencing system and a video lecture.

Results. As data sources we used transcripts of the discourse of the learner as well as learners' text production in the pre-test and in the post-test. In so doing, we identified all scientific concepts explicitly stated or paraphrased in the students' essays in the sequence of their occurrence. Thus, data were simply sequences of concepts, which is a format easily to be analyzed via knowledge tracking. First, we conducted a comparison between results from the pre-test and post-test (analyzed via the expert structure) and the random structure. There was a significant difference between the pre-test and random structure ($z = -1.67$; $p < .05$) as well as the post-test and the random structure ($z = -2.54$, $p < 0.005$) of both groups and the random structure, showing that subjects had prior knowledge that clearly increased during the collaboration. Second, we assessed the degree of convergence that was initiated by either the microscript or the macroscript condition. To assess convergence of knowledge we used the structures derived from traces (generative mode of KT) of subjects of each triad to analyze the traces of subjects of each triad (confirmative mode of KT) avoiding of course circular assessments. In this way, scores for the convergence of knowledge in the microscript and the macroscript condition were determined. A comparison of both script conditions revealed a marginally significant difference in favour of the macroscript condition ($z = -1.47$ $p < .07$).

REFERENCES

- Fischer, F. & Mandl, H. (2001). Facilitating the construction of shared knowledge with graphical representation tools in face-to-face and computer-mediated scenarios. In P. Dillenbourg, A. Eurelings & K. Hakkarainen (Eds.), *European Perspectives on Computer-Supported Collaborative Learning* (pp.230-236). Maastricht: McLuhan Institute.
- Hewson, P. W., Beeth, M. E., & Thorley, N. R. (1998). Teaching for conceptual change. In . G. Tobin & B. J. Fraser (Eds.), *International handbook of science education* (pp. 99-218). Dordrecht, Netherlands: Kluwer.
- Janetzko, D. (2000). Selecting and generating concept structures. In R. Roy (Ed.), *Knowledge management - A micro-level approach*. London: Springer.
- O'Donnell, A. (1996). Effects of explicit incentives on scripted and unscripted cooperation. *Journal of Educational Psychology*, 88 (1), 74-86.
- Reiserer, M., Weinberger, A., Ertl, B., Fischer, F., Mandl, H., & Jahn, K. (in press). *Scripted collaborative learning in text-based computer-mediated communication - A pilot study* (Research report 140). Munich: University of Munich.
- Roschelle, J. (1996). Learning by Collaborating: Convergent Conceptual Change. In T. Koschmann (Ed.), *CSCL: Theory and Practice of an Emerging Paradigm* (pp. 209-248). Mahwah, NJ: Erlbaum.
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24(4), 535-585.