THE CRITICAL DECISION METHOD (CDM): A KNOWLEDGE-MAPPING TECHNIQUE

J.S. Weitzenfeld
J.T. Freeman
T.R. Riedl
G.A. Klein

ABSTRACT In a project to develop a course in C-debugging in Bell Laboratories environments, we elected to employ a knowledge elicitation technique, the Critical Decision Method (CDM), developed by Klein Associates to study decision-making in real-world contexts. Although the method previously had been used to develop training materials, it had not been used for a putative single skill in as wide a range of environments as is found in Bell Laboratories, or in so technical a subject matter area. The method provided a rich picture of the both the task environments and debugging expertise in those specific environments. The purpose of this paper is to share our methodological discoveries in the course of the project.

OVERVIEW OF CDM

Critical Decision Method (CDM) is a variant on the Critical Incident Technique (Flanagan, 1954) that was developed by Klein Associates (Klein, Calderwood & MacGregor, 1989; Crandall, & Calderwood, 1989). Informants are asked to describe in their own words incidents in which their expertise was particularly important. The interviewer reviews the story, using the set of CDM cognitive probes. A timeline is developed, and decision points are located. A typical probe for decision points is "Is there anything that I, coming into this situation new, might have done differently?" When the decision points are established, the interviewer probes for the criteria used for decisions. Typically, interviews are conducted with 7-8 subject matter experts (SMEs) in the same domain, but only one two-hour interview is conducted with each. This contrasts with knowledge elicitation techniques used for the development of expert systems, in which many sessions are conducted with only one or a few SMEs. It reflects the fact that CDM has been used to identify the marginal knowledge that distinguishes experts from non-experts, and not for eliciting an exhaustive catalog of the expert's knowledge base.

CDM has several important advantages as a knowledge elicitation method for training and human factors purposes. It allows the informant to define the range of the problem; it allows the informant to define which problems are most challenging; it permits the informant to identify where experience makes most difference; and the timeline facilitates identification of relatively difficult operations.

SAMPLING

We conducted a preliminary survey of supervisors (Weitzenfeld, Freeman & Riedl, 1989) in the target population for three reasons: To economize in the interviews by finding out in advance about the general task context and availability of resources. With this background information in hand, we could focus the interviewing time on actual debugging efforts. To validate an earlier needs assessment study about the size of the potential audience for the course. To assess the variability in the task domains. The survey suggested that there was a very wide variability both in the size of projects and in the information provided with a debugging task, confirming the need to diversify our sample by project size and geographical location. We had antecedent reasons to suspect both of these variables might affect expertise. The survey raised the possibility that we would elicit environment-specific expertise. Since previous CDM experience had indicated that interviews with 3-4 experts were both necessary and sufficient to grasp the outlines of the relevant knowledge, we conducted at least that number of interviews at each of three sites, including projects of different sizes and both male and female debuggers (Weitzenfeld, 1989). However, if we had found insignificant variation in the first 8 interviews, we were prepared to stop sampling to conserve resources. In the event, we noticed what appeared to be differences in strategy between our first two sites, and so completed our sampling plan.
We also elected to include two informants who had been identified as less expert, or "more representative of the target population," to see if their relative lack of expertise would be apparent to interviewers unfamiliar with the skill. In all, we interviewed 17 debuggers at Indian Hill, Columbus, and Middletown. They worked on projects ranging in size from medium to the large.

THE INTERVIEWING TEAM

Teams comprised of 4 of the 5 team members conducted most of the interviews. The principal interviewer conducted the sessions and contributed expertise in the psychology of problem-solving and decision making, as did the team member who manned our videotape recorder. The project manager doubled in these sessions as a domain expert, translating technical terms and concepts for the interviewer where necessary. The fourth participant was familiar with the empirical literature on program debugging. All team members were psychologists; two had considerable programming experience. We found that some prior technical expertise on the part of the interviewer was far more effective than relying on communication between two people, one expert in CDM and one familiar with debugging and Bell Laboratories. It took a while for the Bell Laboratories personnel who had not had experience with the method to arrive at an optimal level of participation, providing clarifying information but not disrupting the logic of the CDM interview, as well as for the method experts to learn technical material.

INTERVIEW

Our format varied somewhat from interview to interview as we found different approaches useful. The following represents the method we found most effective. The principal interviewer began the sessions by introducing the subject to the team and to the goals of the project. To minimize each subjects' anxiety, the interviewer assured the subject that we would respect proprietary information, and maintain the confidentiality of the interviews and our records of them. The interviewer's explanation of the method was brief: "We'd like to hear your stories about debugging code. We're looking for good detective tales, incidents that challenged you and might have defeated a less experienced professional." Thus, the subjects selected cases and, as described below, isolated the critical incidents in them that they felt demarcated expertise. In CDM, the expert, not the researcher, defines the domain.

Each story was reviewed four times. In the first, the subject described an incident briefly with little interruption from the interviewer. Stories covered the range from shared memory overwrites, to logic errors, protocol violations, and misprogrammed laboratory testing equipment. During this recitation, the interviewer focused primarily on understanding the story. A timeline of events on which the remainder of the interview would be based was created privately by the recorder.

In the second stage, the recorder reproduced his timeline on a whiteboard. The timeline consisted of roughly six to twelve subevents (as determined by the interview), with key facts, assumptions, considerations, relative dates (day 1, day 2, etc.) for each. He also gathered and recorded background information concerning the origin of the problem, information that accompanied the MR, time and political pressures during the period, and the people and systems involved. The recorder encouraged the subject to correct and verify the outline of the event. Consequently, the record on the board, which was elaborated in the subsequent discussion and preserved on the videotape, was a representation of the episode that had been reviewed by the informant. The interviewer spent this time planning the remainder of the interview and noting incidents to probe.

The interviewer then began an in-depth review of the incident. His queries of the subject were opportunistic, but generally included the following for every substage of the incident:

What did you know at this point?
What options did you consider?
Why did you choose this option?
Were there any options you did not consider?
Did this remind you of any previous experience?
What experience or training was needed for this decision, or would have helped?
Who was involved?
What documents did you use?
What equipment did you use?

Answers to these questions provided a rich picture of the problem state as the subject recognized it at each juncture, as well as a rough ranking of available options. The interviewer recorded all of this information, typically filling the whiteboard by the end of the review. Having completed the detailed review, the interviewer asked the subject to identify the deepest pitfalls for the less expert. "If I was new and responsible for this project," the interviewer asked, "what mistakes might I have made?" The interviewer mapped this information, too, onto the timeline, typically using a contrasting pen color. Most subjects were mentors, and had extensive experience with nonexpert behavior.

At the conclusion of the CDM interview, informants were occasionally encouraged to comment and fill us in on issues that had been raised, such as the managerial context for debugging. To wrap up each session, we offered to send our findings to the subject, asked permission to call with follow-up questions, and thanked the subject for contributing to the project.

**CONCEPT MAPPING**

In several of the first interviews, a fifth stage followed. The interviewer erased the timeline and drew a concept map in its place. Each map took the form of a semantic network with objects connected by relations (e.g., "object(audits)", "relation(learned from)", "object(manuals)"). These concept maps were then reviewed by the SME. They proved less illuminating for course development purposes than we had hoped, though one informant remarked on them quite positively.

**DATA GATHERING**

In CDM, typically the first interview serves primarily to orient the interviewer. The quality and detail of the data increase as successive interviews are conducted. Because of the highly technical nature of the information, we feared that the orientation period in the debugging study might be protracted, and so we took great care to preserve the data for subsequent analysis. The interviews were videotaped, notes were taken on the blackboard and later re-copied, and a member of the team (the recorder) created a partial transcript in real-time on a portable computer. These precautions proved useful. The entire team reviewed the videotapes of the first set of interviews, and a second interview with the first informant was held to resolve the questions that arose.

The real-time transcript was a convenient vehicle with which to review the interviews and examine the timelines in meetings shortly following each session. The recorder tested two approaches to note-taking. Summarizing blocks of dialogue increased his ability to contribute, but decreased the usefulness of the notes. As he formulated and typed summaries, he lost ongoing comments. Near-verbatim transcription proved more helpful, even in conjunction with the tapes as a primary data source.

The videotape was the most precise record of each interview, and, accordingly, the most dense. McGraw and Harbison-Briggs (1989) have noted that it can take two to four times as long to review a videotape as to record the original event. Our experience supports this.

In several instances, we had complete transcripts made of the videotapes. While these proved useful, producing them was difficult. First, our transcribers were equipped to work with audio, not video cassettes. Transcribing took longer and was more expensive than we expected. Second, the quality of sound varied on these tapes. The conversational, even intimate, tone, that CDM encourages in informants, made it difficult for transcribers to understand the tapes. Based on these experiences, we recommend using lapel mikes (or high-quality omnidirectional table mikes) mixed through an audio cassette recorder. The recorder output can be piped into the video camera as a backup audio source. Finally, the technical language used by our subjects baffled some transcribers. In the printed interviews, we found some system terms transformed into familiar New Jersey town names.
ANALYSIS

Our first method of analysis was a focus on subjects' decision-making strategies. Our second method of analysis was to scan the tapes or transcripts to sort the material into the following categories: Debugging skills -- such as how to set break points. Knowledge -- such as recent changes in the system. Strategy -- such as choosing to conduct a binary search. Maxims -- such as "Reproduce the problem before reading code." And subjects' suggestions. Our third and most domain-specific method was to review the materials for resources used on the job. Among these were:

Documents
People
Hardware and software tools
Domain knowledge
General contextual knowledge
Social skills
Procedural skills

VERIFICATION

We planned to follow the CDM interviews with a verification stage to test the validity and completeness of our findings. The richness and the depth of insight that debate among experts might provide, and a tight course development schedule lead us to verify our findings with a panel discussion among some of our informants.

ADVANTAGES

A significant advantage of CDM is its psychological validity. CDM is predicated on expert memory for stories. This is supported by the detail available in episodic memory (Tulving, 1972) and the advantages of recall with structured probes. A second advantage of CDM derives from its character as an open-ended process. The expert subject, not the researcher, defines the domain by choosing events and critical incidents within them. CDM is, thus, a particularly useful strategy in complex, novel, or real-life domains, where the researchers are relatively naive. In virtually every empirical study of debugging to date (Curtis, et al, 1986; Sheils, 1981; Brooks, 1980) the research paradigm has been to give individual subjects novel code to debug. Every incident of debugging discussed by our subjects involved teams of professionals. Virtually all concerned code with which the subjects were familiar. Allowing subjects to define the environment of debugging was an invaluable aspect of the CDM method in this project. CDM also offers several purely practical benefits. It uses expert time efficiently and economically. Sessions consume only a few SME hours. Unlike card-sorting or repertory grid techniques, CDM also appears to be enjoyable to experts. Most are flattered to tell their stories, and the story-telling mode seems to dissipate social tension. Finally, the interviews serve both as data and product. We not only analyzed the videotapes, but hope to secure subjects' permission to use video clips and transcript excerpts in training.

LIMITATIONS

Certain aspects of our informants' expertise surfaced at the follow-up conference. Our inability to clarify these aspects from the interview data shed light on some weaknesses of CDM. First, we detected cross-site differences among our informants that were difficult to analyze. We first inferred a difference in strategy, and then the presence of a strategic rule -- (roughly) "Do cheap searches first" -- that never functioned in environments in which no search was cheap. This situation was clarified at our conference. It turned out that there were gross differences among the various sites in the accessibility of resources. Each informant worked primarily in one environment, so, although their expertise had been shaped by the environment, SMEs never made an overt choice of strategies based on resource availability. Some options were so difficult or time-consuming to execute at certain sites that they were simply not considered, and they did not surface as decision points in the CDM interviews. The moral here reinforces the importance of sample size and selection informants, as well as paying specific attention to factors in the task environments that are beyond their control. Note that the behavioral differences were noticeable from the CDM interviews, but their cause was not. The second major factor to surface at the follow-up conference was a non-linear
problem-solving mode, a periodic reviewing of all the information that often lead to insight. CDM elicited a rational search process, with a role played by recognition-primed decision (Klein, 1989). Again, the interviews had provided some clues that something was missing, but the panel session clarified somewhat what it was.

**USES OF CDM**

In addition to descriptions of expertise that could be used goals for training or recruiting, our interviews produced data about organizational practices, the kinds of faults that consumed large amounts of debugging time, the effects of resource availability, and the work climate. Moreover, it brought these topics to our attention when we may not have thought of asking about them. This suggests that the method can be used for organizational design and effectiveness studies, recruiting, human factors, and process quality improvement, as well as for education. It also suggests that this or similar methods should be used on a regular basis to monitor the quality of the work environment.

**REFERENCES**


