DNA - Uncorking the Bottleneck in Knowledge Elicitation and Organization

Valerie J. Shute

Air Force Research Laboratory (AFRL)
1880 Carswell Avenue
Lackland AFB, TX 78236-5508
vshute@colab.brooks.af.mil

Lisa A. Torreano

National Research Council Associate
1880 Carswell Avenue
Lackland AFB, TX 78236-5507
lisat@colab.brooks.af.mil

Ross E. Willis

Galaxy Scientific, Corp.
1880 Carswell Avenue
Lackland AFB, TX 78236-5507
rwillis@colab.brooks.af.mil

Abstract. There are two main purposes of this paper. First, we describe a novel cognitive tool that was designed to aid in knowledge elicitation and organization for instructional purposes – specifically to be used for intelligent tutoring system development. This automated approach to knowledge elicitation is embodied in a program called DNA (Decompose, Network, Assess). Our aim for this tool is to increase the efficiency of developing the expert model – often referred to as the bottleneck in developing intelligent instructional systems. The second purpose is to present a first-order summative evaluation of the tool's efficacy. Specifically, we used DNA with three statistical experts to explicate their knowledge structures related to measures of central tendency. In short, we found that DNA can be used as a standalone program to effectively elicit relevant information on which to build instruction. This was achieved in hours compared to months for conventional elicitation procedures.

1   Introduction

One of the most significant challenges we face in the prevailing "information age" is knowledge management. For information to be useful, it must be easily accessible and sufficiently organized so that it both makes sense and can potentially be conveyed to others. This challenge, however, is by no means new – especially to those persons in the teaching profession. Anyone attempting to design instruction must first develop a sound curriculum. In all cases, *what* information to include in the curriculum and *how* to ensure learners' mastery of the material must be determined, regardless of whether instructing karate beginners, nuclear physicists, mechanics, or artists.

These issues are particularly salient in computer-assisted instruction. To render such systems intelligent, and hence, more effective and efficient, three components must be included in the system: (a) an expert model, (b) a student model, and (c) an instructor model [1], [2]. The expert model represents the material to be instructed. This includes both elements of knowledge related to the domain to be instructed and their associated structure or interdependencies. In essence, the expert model is a knowledge map of what is to be taught. The student model represents the student's knowledge and progress in relation to the knowledge map or curriculum. The instructor model (also known as the "tutor") manages the course of instructional material and remediation strategy based on discrepancies between the student and expert models. Thus, the instructor model determines how to ensure learner mastery by monitoring the student model in relation to the expert model. The strength of these programs, therefore, depends on the underlying validity of the curriculum or domain expertise [3].

There are two main purposes of this paper. First, we describe a novel cognitive tool that we have designed to aid in knowledge elicitation and organization for instructional purposes – specifically to be used for intelligent tutoring system (ITS) development. This automated approach to knowledge elicitation is embodied in a program called DNA (Decompose, Network, Assess). Our aim for this tool is to increase the efficiency of developing the expert model – aptly referred to as the backbone of intelligent instructional systems [3]. We will summarize its interface and functionality, but refer the reader to more detailed information on the program [4].

The second purpose of this paper is to present a first-order summative evaluation of the tool's efficacy. We outline the results from an empirical validation of the tool that examined how efficiently and effectively DNA works in the extraction of knowledge elements related to statistics. Specifically, we used DNA with three statistical experts to explicate their knowledge structures related to measures of central tendency.

To satisfy this paper's goals and put them in perspective, we will begin with a short summary of current methods of knowledge elicitation and organization. Next, we will describe the DNA program, illustrating the niche our particular tool fills in the intelligent tutoring system arena. This will be followed by a discussion of the first-order summative evaluation results that enables us to assess the potential utility and effectiveness of DNA.

1.1   Techniques of knowledge elicitation and organization

A variety of cognitive task analysis (CTA) techniques are united by the common goal of representing the underlying knowledge, skills, and goal structures of a particular domain or task performance. Typically, CTA is done in two phases: (a) initial acquisition of knowledge and skills, and (b) compilation of information into a useful database that can serve many functions (e.g., curriculum for instruction or an expert system).

How is information initially acquired? Acquisition of knowledge and skills from experts can take the form of observations, verbalizations, and interviews. The most unobtrusive procedure for obtaining data from experts involves monitoring or observing them as they perform their job or specific task. Another technique, "think aloud" protocols, involve having experts verbalize their thoughts as they perform some task, answer questions, or solve problems. Finally, interviews are a common method for eliciting expert knowledge and skills. Typically, the series of questions posed to the expert are designed around a theory of expertise or the framework in which the information will subsequently be used, e.g., in designing a cognitive model or instruction [5]. All of these approaches can vary in a number of methodological ways including whether the knowledge elicitation is structured or unstructured, and whether it is retrospective or concurrent with task performance. One drawback to these methods is that people do not always have accurate access to their cognitive processes [6]. Thus, these methods may be weak in delineating certain underlying cognitive operations or in tapping into knowledge and procedures that have been automated.

How is information optimally represented? Conceptual graphs are one popular means of representing hierarchically-structured knowledge. As the name implies, conceptual graphs are the graphical representation of concepts showing, at various grain sizes, relevant concepts (nodes) and their interrelationships (arcs or links). Another popular representation of knowledge, a means-end hierarchy, results from a production system framework using a GOMS-type analysis (GOMS= Goals, Operators, Methods and Selection rules) [7]. Each representational type produces a knowledge structure that has its own unique flavor, e.g., conceptual vs. procedural in the two examples, cited above.

Regardless of the approach for eliciting and/or structuring knowledge, current CTA procedures are usually very labor-intensive, often not standardized, potentially incomplete, and difficult to translate into instruction and assessment. For good reviews of CTA methods, see Shraagen [8] and Williams [9]. Overall, CTA remains a bottleneck in the ITS development process.

1.2   Uncorking the Bottleneck: Description of DNA

We are designing a novel cognitive tool called DNA (Decompose, Network, Assess) to help alleviate this bottleneck. Automating the bulk of the cognitive task analysis procedure can potentially expedite the CTA process without sacrificing

accuracy. In addition, our goal is to create a cognitive tool that is applicable for systematic decomposition of any domain into its constituent knowledge and skill elements, whether related to task performance (e.g., troubleshooting jet engines) or not (e.g., understanding the core concepts of world religion) [4]. DNA is used to Decompose a domain, Network the knowledge into comprehensive structures, and employ other experts in a given domain to Assess the validity, completeness, and reliability of the knowledge structures. This approach is embodied within an easy-to-use program aimed at extracting and organizing knowledge and skills from experts.

1.3   Modules of DNA

Customize. The Customize module is completed by the instructional designer (ID) who provides information about the ultimate learning goal of the tutor to be developed, the prerequisite knowledge and skills of the learners, and the desired instructional emphasis or flavor (e.g., primarily procedural). This information provides the SME with the superordinate goal of the analysis and the lowest-level subordinate goal at which point the SME should stop decomposing the domain. Using the instructional designer's input, the Customize module generates a personalized letter that explains the purpose of the project and a set of floppy diskettes, which are to be mailed to prospective SMEs. The diskettes contain files for a SME to install on his or her computer that DNA needs to elicit and store knowledge structures.

Decompose. The Decompose module is a semi-structured, interactive dialog between computer and SME designed to elicit most of the explicit knowledge associated with the domain/topic of analysis. DNA utilizes the "What, How, Why" questioning procedure which has been shown in the past to successfully elicit knowledge from experts (e.g., [10]). "What" questions (e.g., "What is a definition of _____?") are designed to obtain symbolic knowledge, or definitions of terms. "How" questions (e.g., "How do you _____?" or "What is the first step you do when you _____?") are designed to elicit procedural knowledge and skills. Finally, "why" questions (e.g., "Why is _____ important?") are designed to elicit conceptual knowledge, or higher-level relationships among concepts.

   All information given by experts is stored in a database as a record of curriculum elements (CEs) which are components of knowledge needed to develop expertise in the domain. By storing information in a CE record, it should become easy to translate the information into teachable units and thereby generate curricula.

Network. The Network module transforms CEs elicited during the Decompose module into graphical nodes which experts spatially arrange and link to form knowledge hierarchies, conceptual graphs, or production rules. Each node contains the name of the CE and its contents as defined during the Decompose module. Links placed between CEs can differ in terms of *strength* (i.e., weak, moderate, strong—showing the degree to which the items are related), *type* (e.g., is a, causes, fixed serial order, etc.—denoting the nature of the relationship between nodes), and *directionality*

(i.e., uni-, or bi-directional—indicating which CEs are prerequisites to other CEs and the flow of causation).

The use of a graphical representation should make relationships among knowledge units salient, which can also highlight missing knowledge components. This module is similar to conceptual graph analysis except that, with DNA, experts generate the conceptual graphs instead of the instructional designers. Thus, we speculate that DNA will enable experts to recognize gaps in the knowledge and skills they provided earlier. Moreover, they have a chance to readily correct inadequacies as they can return to the Decompose module and update the CE record with new information.

After SMEs complete the Network module, data are stored on floppy diskettes and returned to the ID. The ID reviews the CE record and conceptual graphs for any glaring omissions in content. If any omissions are present, the ID can ask the expert to expand the inadequate CEs.

Assess. The final module is used to validate the CE records and conceptual graphs generated by SMEs. This is accomplished by having other experts in a domain review the conceptual graphs generated by the first group of experts. Multiple experts can also be used to edit initial conceptual graphs as a method of modifying and validating the externalized knowledge structures.

The modules are repeated until the instructional designer is satisfied with the content of the revised CE records and conceptual graphs. Efforts have been made to elicit a range of outcome types: symbolic knowledge (SK), procedural knowledge and skills (PK and PS), and conceptual knowledge (CK).

1.4   DNA Promises and Purposes

Conventional CTA methods have been described as being inefficient, limited to procedural tasks, laborious to translate into usable instruction, and difficult to use. We have attempted to address each of these limitations when we designed DNA [8].

Efficiency. Traditional CTA methods typically involve extensive interviews with experts, transcription of ensuing protocols, and organization of knowledge and skill units. This process normally requires many months of work and many person-hours to achieve. These traditional methods often employ knowledge engineers to interview and observe experts, others to transcribe sessions, and cognitive psychologists to summarize the data into a hierarchical representation. In contrast, DNA automates the bulk of the interview, transcription, and organization processes which is intended to significantly decrease both time and personnel resources required for the cognitive task analysis.

Broad applicability. A common limitation of traditional CTA methods is that most are only applicable to procedural domains. DNA's specific purpose, and thus strength, relates to ITS development and being applicable across a range of domain topics - procedural and conceptual in nature. This is achieved via its underlying

hybrid representational structure of knowledge and skills elements—functionally a cross between a semantic net and production system.

Instructional design framework. Another common limitation of traditional CTA methods is that it is often difficult to translate the pages of interview transcriptions and conceptual graphs into a usable curriculum. DNA is structured such that its output is directly compatible with an existing instructional system framework (i.e., SMART; [11]) which further enables efficient tutor development. DNA's resulting database of CEs contains rich and useful information at the individual CE level (e.g., unique CE number, outcome type category, detailed description, and hierarchical information such as parents, siblings, and children). The database also includes embellishments from the SME such as: typical points of difficulty (impasses) in understanding CEs, good examples and counter-examples of CEs, and specific questions that can serve to assess understanding of CEs. All of this information is well suited for developing subsequent and principled instruction.

User-friendly. As previously mentioned, traditional CTA methods often rely on several individuals trained in knowledge elicitation techniques. However, DNA was designed to be very user-friendly, thus the program is not limited for use by individuals with CTA experience. Rather, any instructional designer who wants to develop a curriculum will be able to use this tool to elicit knowledge.

2   Preliminary DNA Evaluation

What we present here is the summary of an investigation testing the Decompose module of the DNA system. In general, the different evaluation issues relate to the system's consistency, validity, and efficiency. DNA promises a great deal in its potential to render the CTA process faster, cheaper, and better. However, before assessing its relative benefits, we need to determine a more fundamental issue: Can DNA indeed extract comprehensive knowledge from subject-matter experts that could serve as the basis for curriculum development?

2.1   Participants and Design

Three volunteer subject-matter experts participated in this preliminary study. While none are formally "statisticians," all have graduate degrees in psychology and a minimum of 10 years experience conducting statistical analyses. Further, all reported that they were quite familiar with "measures of central tendency."

Before the experts arrived, the authors completed the Customize module of DNA to produce the letter informing the experts of the curriculum goals for students to achieve. These included: (a) identify the main measures of central tendency, (b) identify the formulas for the measures of central tendency, (c) know how to calculate each measure of central tendency, and (d) understand the relationship(s) between each

measure of central tendency and different underlying distributions. In addition, this letter informed the SMEs that the intended learner population would have (a) basic math skills (with algebra), (b) basic reading skills, and (c) basic computer skills (Windows 95 interface). This provided the SMEs with parameters for their decomposition of the domain. Each SME devoted several hours to decomposition via DNA, and at least one of the authors was present during each of the three interactions to answer only general questions posed by the SMEs.

We are interested in testing the degree to which our SMEs' data agree with a benchmark representation of the domain. Williams [9] conducted a similar analysis using a production system representing cutting and pasting text with a word processor. We are interested in seeing whether this technique can be used for more conceptual data as well. Therefore, because we have an existing tutor that focuses on this topic (i.e., Stat Lady; [11]), we chose the domain of "measures of central tendency". The curriculum for this module of Stat Lady was derived from a traditional cognitive task analysis involving document analysis and interviews with SMEs. The CEs that were obtained for the Stat Lady curriculum required about six months to obtain.

The degree to which knowledge elements derived from the experts using DNA map onto a curriculum already embodied in an existing tutor, will shed light on the consistency of DNA's output as well as potential effectiveness, a validity issue. That is, if the obtained outputs are close to the "idealized" domain structure of a tutor that has already been shown to improve learning, we can infer that the outputs are valid.

To assess incoming levels of expertise, the SMEs completed an on-line test of measures of central tendency used in conjunction with Stat Lady. The test assesses knowledge and skills related to all CEs contained within the Stat Lady curriculum (a total of 127 CEs). Our experts required between 1-1.5 hours to complete the test (no time limits were imposed), and scores ranged from 71.3% to 87.5%. Ideally, however, expert scores would have been in the 90% range. Following the test, each expert completed the Decompose portion of DNA.

2.2   Analysis and Results

The output from DNA comes in two forms: (a) Microsoft Access database of CE records, and (b) graphical array of the hierarchical knowledge structure. However, because the focus of this DNA evaluation was on the Decompose module, only the CE databases were analyzed for the consistency assessment.

We used the Stat Lady database of curriculum elements as the basis for comparing completeness and consistency of output among the SMEs. A few example CEs from Stat Lady include knowing: the notation for sample size (N) and summation (Σ), the steps involved in the computation of the mean when all frequencies are 1, and the location of the mean within a negatively-skewed distribution relative to the median and mode. In total, 78 CEs from this particular Stat Lady module were appropriate for the topic of analysis (i.e., measures of central tendency).

The first stage of analysis involved assessing the contents of each SME database. We created a spreadsheet that included a column listing all Stat Lady CEs, as well as three columns representing our SMEs' data. We then placed a "1" next to the Stat Lady CE to indicate that the SME did, in fact, include it in his or her decomposition and a "0" to denote its absence. Partial credit was also assigned for some CEs. For example, if there was a 3-step procedure, and an expert only reported 2 of the 3 steps, we assigned a value of .67 for that CE. And finally, there were some instances where a SME delineated a CE that was not present in the benchmark listing. Those instances were noted, but not included in the current analysis.

How well do the experts capture the benchmark curriculum? Our three SMEs matched the Stat Lady database 25%, 49% and 23% respectively. Further, each required 285, 170, and 100 minutes to complete DNA. One expert (SME-2) was clearly more in line with Stat Lady than the others, producing her array of CEs in less than 3 hours of decomposition time. The outputs produced by all three experts were subsequently combined to determine the degree of total overlap with the benchmark. Results showed that 62% of the Stat Lady CEs were delineated by at least one of our three experts. Thus, the agreement between the aggregate and benchmark data showed that we can capture 62% of the CEs present in an existing database in a reasonable amount of time (i.e., approximately 9 hours, the total of all time required by the 3 experts).

With regard to our benchmark CEs, some of them were reported by more than one expert, while others were omitted completely. The duplicated CEs included definitions of the Mean, Median, and Mode, as well as the basic steps required in determining the values of each of these measures of central tendency. Other knowledge bits specified by multiple experts related to information about distribution types and relevant notations. Benchmark elements that were omitted by all three experts included the specification of: (a) particular types of distributions (e.g., leptokurtic), (b) the location of the central tendency measures within distributions, and (c) different ways to compute measures of central tendency (e.g., finding the Median when N is odd versus even).

These data provide preliminary information about the efficacy of DNA as a knowledge elicitation tool. That is, given limited direction via one introductory letter of expectations for the decomposition of the domain and minimal guidance in use of the DNA program, experts appear to be able to use the tool to efficiently explicate their knowledge structures. Moreover, the obtained data are consistent with an existing curriculum. Thus we are gaining confidence that our tool has potential value as an aid to the bottleneck in ITS development.

3   Summary and Discussion

This paper describes an ongoing effort to develop a knowledge-elicitation tool called DNA, to be used by subject-matter experts across a variety of domains. We also describe an exploratory test of the effectiveness and efficiency of the program. Preliminary results show that DNA can produce valid and reliable data within

reasonable amounts of time. This has direct implications for streamlining the ITS development process, previously defined as the bottleneck in developing automated instructional systems. In addition, given these data were obtained from individuals who are not "statisticians" suggests that DNA can be used by persons varying in levels of expertise.

There are several key features of DNA that, we believe, make this a viable alternative to current and costly knowledge-elicitation techniques. That is, it is automated, uses a hybrid representational scheme, provides an interplay between elicitation and organization of evolving knowledge structures, and is based on an effective instructional design framework (SMART; [12]).

In general, DNA is a computer program that automates the interview, transcription, and organization processes. This automation allowed us to obtain data simply by giving each expert the program along with a short letter explaining the goals of the curriculum. Thus automation obviates the need for transcribing lengthy interviews. Additionally, experts are able to explicate and organize their knowledge within the same elicitation session, translating into expected savings of time and money without sacrificing accuracy. This will be examined in future studies.

DNA's applicability is enhanced because it elicits, and then allows SMEs to graphically represent, a range of knowledge types. Specifically, the Network module can produce a conceptual graph that is a hybrid between a production-rule (or GOMS-type) structure and a semantic net. The result is that this hybrid representational scheme enables DNA to obtain both procedural- and conceptual-flavored curricula, promoting applicability across multiple topics.

Another design feature of DNA is its compatibility with an empirically validated instructional framework (i.e., SMART; [12]). First, DNA's Network module empowers the SME with tools to create a complete hierarchical structuring. The SME can create and manipulate CEs, as well as establish links between them and specify their relationships. An accurate representation of the knowledge structure directly affects the efficacy of the expert model used in an ITS, and thus learning. Specifically, SMART relies on information present in hierarchical-knowledge structures (e.g., parent/child relations) to manage instruction and remediation. Second, DNA's Decompose module's what, how, and why questions map onto the instructional framework of symbolic, procedural, and conceptual knowledge types embodied by SMART. Specifically, SMART relies on these knowledge types to provide differential instruction and remediation. For instance, procedural knowledge is instructed within a problem-solving context, while conceptual knowledge may use analogies for instruction. Therefore, DNA's capacity to identify different knowledge types facilitates SMART's management of more customized instruction.

In conclusion, the results from this preliminary evaluation are encouraging. In a relatively short amount of time and with minimal resource cost, DNA (the Decompose module) was able to elicit more than 60% of the curricular elements that are in place in an extant tutor. Our initial question for this paper concerned whether, indeed, DNA can extract comprehensive and reasonable knowledge from experts. Results suggest that the general approach implemented by DNA works to produce valid data that could potentially serve as the basis for curriculum development. Future

studies will examine DNA's efficiency relative to standard knowledge elicitation techniques. Additional questions we will explore include, among others: (a) Can DNA be used across a broad range of domains? (b) Is it differentially effective in eliciting symbolic, procedural, or conceptual knowledge elements? and (c) Do differing levels of expertise result in data structures that vary in kind, rather than simply quantity? In short, future development and research will focus on identifying where we have and have not succeeded in our aim to uncork one of the main bottlenecks in ITS development.

References

1. Polson, M. C., & Richardson, J. J.: Foundations of Intelligent Tutoring Systems. Lawrence Erlbaum Associates, New Jersey (1988)
2. Shute, V. J., & Psotka, J.: Intelligent Tutoring Systems: Past, Present, and Future. In: Jonassen, D. (ed.): Handbook of Research on Educational Communications and Technology: A project of the Association for Educational Communication. Macmillan, New York (1996) 570-600
3. Anderson, J. R.: The Expert Module. In: Polson, M. C., Richardson, J. J. (eds.): Foundations of Intelligent Tutoring Systems. Lawrence Erlbaum Associates New Jersey (1988) 21-50
4. Shute, V. J., Torreano, L. A., & Willis, R. E.: Cognitive Task Analysis: Towards an Automated, Powerful Tool. In: Lajoie, S. (ed.): Computers as Cognitive Tools, Vol. 2. Lawrence Erlbaum Associates, New Jersey (in press)
5. Ryder, J. M., & Redding, R. E.: Integrating Cognitive Task Analysis Into Instructional Systems Development. Educational Technology Research and Development 41 (1993) 75-96
6. Nisbett, R. E., & Wilson, T. D.: Telling More Than We Can Know: Verbal Reports on Mental Processes. Psychological Review 8 (1977) 231-259
7. Card, S., Moran, T. P., & Newell, A.: The Psychology of Human-Computer Interaction. Lawrence Erlbaum Associates New Jersey (1983)
8. Shraagen, J. M. C., Chipman, S. E., Shute, V. J., Annett, J., Strub, M., Sheppard, C., Ruisseau, J. Y., & Graff, N.: State-of-the-Art Review of Cognitive Task Analysis Techniques. Deliverable Report of RSG.27 on Cognitive Task Analysis NATO Defense Research Group (Panel 8/RSG.27). TNO Human Factors Research Institute Group: Information Processing (1997)
9. Williams, K. E.: The Development of an Automated Cognitive Task Analysis and Modeling Process for Intelligent Tutoring System Development. Contract final report on N00014-97-J-5-1500. Manpower Personnel and Training Program, Office of Naval Research (1993)
10. Gordon, S. E., Schmierer, K. A., & Gill, R. T.: Conceptual Graph Analysis: Knowledge Acquisition For Instructional System Design. Human Factors 35 (1993) 459-481
11. Shute, V. J., Gawlick, L. A., & Lefort, N. K.: Stat Lady [Unpublished computer program]. Brooks Air Force Base, TX: Armstrong Laboratory (1996)
12. Shute, V. J.: SMART: Student modeling approach for responsive tutoring. User Modeling and User-Adapted Interaction 5 (1995) 1-44