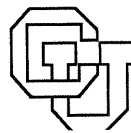


**Using the Interface to Improve Performance
of Complex Cognitive Tasks**

D. Charles Hair

CU-CS-490-90



University of Colorado at Boulder

DEPARTMENT OF COMPUTER SCIENCE

**ANY OPINIONS, FINDINGS, AND CONCLUSIONS OR RECOMMENDATIONS
EXPRESSED IN THIS PUBLICATION ARE THOSE OF THE AUTHOR(S) AND
DO NOT NECESSARILY REFLECT THE VIEWS OF THE AGENCIES NAMED
IN THE ACKNOWLEDGMENTS SECTION.**

Using the Interface to Improve Performance
of Complex Cognitive Tasks

D. Charles Hair

CU-CS-490-90

October

Institute of Cognitive Science and
Department of Computer Science
Campus Box 430
University of Colorado
Boulder, CO 80309
303-492-4932
hair@boulder.colorado.edu

Using the Interface to Improve Performance of Complex Cognitive Tasks

D. Charles Hair
Institute of Cognitive Science
Computer Science Department
Campus Box 430
University of Colorado
Boulder, CO 80309
303-492-4932
hair@boulder.colorado.edu

Introduction

This paper describes work involving the idea that computer interfaces can be constructed that enhance the performance of humans at complex cognitive tasks. The specific idea being investigated is that a hypertext based graphical interface can be tailored to improve people's ability to construct legal arguments. To that end, a tool called LEGALESE has been built. To test the effectiveness of LEGALESE experiments have been run in which LEGALESE is used by subjects from the legal field to aid in constructing legal arguments. Those arguments have been compared to arguments prepared without the aid of LEGALESE.

Background

The idea of using a hypertext-like computer interface to improve people's performance of complex cognitive tasks arose with the pioneering ideas proposed by Bush and Engelbart [1,4]. Engelbart in particular has adhered strongly to the idea that hypertext can be used to directly improve performance at a variety of intellectual tasks, including argumentation.

The specific area of argumentation has attracted a number of other researchers interested in developing hypertext based systems to support the argumentation process. The first such project was called IBIS [3,8], and that project has spawned some related projects being pursued currently at MCC [3] and at the University of Colorado [8]. A project directly related to legal arguments is under way at Xerox PARC [7]. A third argumentation project, called EUCLID, is being pursued at the University of Colorado [9].

The LEGALESE work began as an outgrowth of the EUCLID work. It was decided that the legal field would be a natural one in which to test ideas about computerized argumentation tools. Moreover, as a lawyer myself I have been able to draw on my own expertise, as well as on my ability to communicate with other lawyers in their own terms while discussing this work with them. The emphasis of the LEGALESE work is on exploring the specific question of whether a computerized argumentation tool can be used to help people create better arguments.

The LEGALESE Model of Legal Argumentation

The LEGALESE model is quite similar to the IBIS model of arguments. The basic structure of a LEGALESE argument is to have a hierarchical framework of issues, which are supported by arguments, which are in turn supported by facts and laws. The term *argument* is used here in two ways. At the outer level *argument* refers to the overall structure of what can be an arbitrarily complex approach to a legal question. At the internal level, *argument* refers to

separate pieces of the larger structure that are identified as arguing for a particular resolution of a particular issue.

Because of the required hierarchy, LEGALESE imposes the requirement that a user always begin constructing an argument by stating a single main issue. All subsequent issues are then required to be subissues of existing issues.

Unlike issues, LEGALESE arguments, facts, and laws can be created that are not immediately connected into the overall argument structure. This flexibility is allowed as a result of comments made by a group of lawyers who participated in the design phase of this project [5]. It was pointed out that users sometimes find it desirable to be able to enter their ideas quickly, without going through the process of fitting the idea into the main argument immediately.

One of the basic thrusts of the LEGALESE model is to help the user to focus on identifying the issues in a legal situation. Thus, there must be at least one issue in a structure, all issues other than the main issue must be connected into the overall structure as subissues, and all other parts of the argument should eventually be connected to the issues. The reason for this emphasis on issues is that the ability to identify relevant issues in a legal case, and to tie those issues together in a coherent structure, is at the heart of the correct way to approach constructing legal arguments. It is my hypothesis that the LEGALESE tool will successfully influence users to construct arguments with an emphasis on using the model suggested by the tool. Thus, it is predicted that users will construct arguments such that there is a clear reliance on answering issues with the use of arguments, using facts and laws to directly support the arguments. As a result of using the LEGALESE model of arguments, it is expected that the resulting written arguments will be judged by an independent group of lawyers to be stronger than arguments produced by using other models.

Method

The LEGALESE tool was constructed using HyperCard on a Macintosh computer. Experiments were run on 23 subjects. Of those subjects, 14 were law school students and 9 were practicing lawyers. The lawyers participated on an unpaid basis, as did 3 of the students. The other 11 students were paid for their participation.

In order to test the effectiveness of the LEGALESE tool a hypothetical legal situation was constructed and presented to subjects (see figure 1). The hypothetical situation was written with the idea of being fairly simple, and with the idea of not involving specific legal issues anyone would be likely to have encountered either in a law school class or in the actual practice of law.

Each subject participated on an individual basis. For the experiments, the subjects were placed into one of two groups. An effort was made to make the placement on an alternating basis as subjects became available.

The experiment was divided into two basic phases. In the first phase the subjects in one group (the writing group) were asked to spend up to 30 minutes analyzing the hypothetical situation and making whatever pen and paper outlines they desired. Subjects in the other group (the LEGALESE group) spent an initial 10 minutes being introduced to the LEGALESE tool, primarily through exploration of a prepared example. In their first phase they were

then given up to 30 minutes to analyze the hypothetical situation and construct a LEGALESE representation of their approach to the argument. I was present at all experiments, and offered assistance to the subjects in the LEGALESE group in the use of the tool during the first phase. In the second phase of the experiment, subjects in both groups were given up to 20 minutes to hand write their final arguments, making what use they wished to of the prior outlining efforts or LEGALESE representations.

The lawyers were treated somewhat differently than the student subjects in that the lawyers were allowed to choose to be in the LEGALESE group if they wished. As a result, 8 of the 9 lawyers chose to be in the LEGALESE group. The effort to strictly alternate the subjects was effected by this result, as it became necessary to ask several students in a row to be in the written group in order to balance the groups. Nonetheless, by the time all of the subjects were finished, there were 7 students in the LEGALESE group and 7 in the writing group. The results of the experiment are discussed below both in terms of strictly looking at the student results and in terms of looking at the results when students and lawyers are combined.

The first three subjects were law school students, two of whom were in the LEGALESE group and one of whom was in the writing group. At that point in running the experiments LEGALESE subjects were asked to think aloud [6] while using LEGALESE, and voice recordings were made during the first phase of the experiment. No recording was done of the writing subject. All subsequent subjects from both groups were asked to think aloud during the first phase of the experiment, and videotapes with sound were made of either the process of constructing written outlines or the process of using LEGALESE.

Four of the lawyers did not complete the experiment. At the end, there were 19 written arguments that the subjects had produced, including one lawyer who produced the argument with a dictating machine rather than with pen and paper. Only the data from those 19 subjects who produced the written arguments has been used in the statistical comparisons discussed below. Those arguments were reproduced in typed form, with original spellings and spatial layouts preserved. The arguments were identified only by number, and were sent to an independent group of individual lawyers who agreed to rank the arguments from best to worst.

Results

During the first phase of the experiment, subjects in the LEGALESE group spent a good part of their time in efforts to understand how the tool worked. This extra time was evidenced by repeated questions as to how to accomplish a given task or as to what they should do next. It was also evidenced by the fact that all of the LEGALESE subjects found it necessary to use the entire 30 minutes they had in the first phase, while only two of the writing subjects used the whole time period in writing outlines.

The amount of output produced by both groups during the first phase was approximately equal in terms of the number of words. The LEGALESE group wrote a mean of 189 words while the writing group wrote a mean of 180 words.

In creating their arguments subjects in the LEGALESE group sometimes created argument pieces in which they did not enter any text. Counting such blank pieces, they produced a mean of 8.3 argument pieces. The mean was 7.8

pieces if blank ones are not counted. The overall breakdown of the pieces is: 32 issues (2 blank), 29 arguments (1 blank), 13 cases, 3 statutes, and 13 facts (2 blank).

It is possible in LEGALESE to create argument pieces that are not connected to issues, and to create fact or law pieces that are not connected to arguments. Overall, 16% of the argument pieces were left unconnected (6 arguments, 2 laws, and 7 facts). Where there are no arguments attached to an issue the issue is considered unsupported, and where there are no facts or laws attached to an argument the argument is considered unsupported. Overall, 43% of the argument pieces were unsupported (19 issues and 20 arguments).

During the first phase, the LEGALESE group used complete sentences more often than the writing group. All of the LEGALESE subjects used complete sentences, with two of them using some sentence fragments. All of the writing subjects used sentence fragments, with three of them using some complete sentences. In using the chi square test to calculate whether there was a significant difference in the use of complete sentences between the two groups, I counted only the 9 subjects from the LEGALESE group who only used complete sentences, and only the 3 subjects from the writing group who used some complete sentences. There was a significant difference at the .1 level.

It appeared that three of the LEGALESE subjects did not make appropriate use of the LEGALESE tool. One of them produced a lot of prose (340 words) inserted into only 5 argument pieces. That was the highest word count of any subject in either group. Another subject indicated throughout the first phase of the experiment a general failure to understand how the tool worked, and a general dislike for using computers. The result was to produce the least amount of prose (33 words) of any subject in either group. A third subject did not indicate a dislike for using computers, but did clearly indicate problems in understanding how to use the tool and also produced very little prose (73 words). The first two subjects were students, and the third was a lawyer.

Rankings of the written arguments were obtained from five lawyers. Those rankings have been averaged to show the ranks given in table 1. Use of the Mann-Whitney U test on all subjects, except for the three LEGALESE subjects discussed above, shows a significant difference at the .1 level between the ranks achieved by the writing group versus the LEGALESE group. There is also a significant difference at the .1 level when the calculation is done only counting the students from the writing and LEGALESE groups, and again not counting the two student LEGALESE subjects discussed above.

A separate analysis was done to determine which of the arguments seemed to adhere most closely to the LEGALESE model of legal argument. I identified 10 written arguments that reflected the approach of focusing on issues, tying arguments into those issues, and tying the laws and facts into the arguments. This analysis was based on constructing diagrammatic representations of the arguments. Eight of the ten were produced by the LEGALESE group. The difference between the two groups based on this characteristic was significant at the .1 level (chi square).

Of the 10 subjects who were judged to have used the LEGALESE approach, 8 were ranked in the top 10 of the overall rankings the lawyers produced. The difference between subjects who used the approach being in the top 10, and

subjects who did not use the approach being in the top 10 was a significant one at the .01 level (chi square).

Comparative statistics using the Spearman rank correlation were calculated to determine the consistency of the lawyers doing the rankings. Only two of the rankings were found to be correlated at a significant level. This result appears to reflect the fact that the bulk of the arguments in the middle ranks were of such close quality that there was difficulty making accurate distinctions between them. The most uniformly rated subjects were the top three and the bottom 6.

Discussion

The fact that by far the largest number of argument pieces constructed were issues and arguments suggests that the LEGALESE subjects did focus more on the parts of the argument that are most important in the LEGALESE model of legal argument. This result suggests that the LEGALESE tool is successful at the level of focusing users' attention where desired. The main drawback suggested is that the tool may have the undesirable effect of forcing users to commit to structured ideas too early in the process of developing their ideas. Thus, there has been discussion in the literature about the desirability of allowing early efforts to proceed in a more unstructured way [2].

The finding that use of the LEGALESE tool does influence the way users construct their arguments is further borne out by the finding that the LEGALESE users were more likely to write arguments that emphasized the LEGALESE approach. There is definite evidence here that the interface has a real impact on the way users approach the task of constructing a legal argument.

The main result suggested by the Mann-Whitney calculations is that there is a clear suggestion that the LEGALESE group did outperform the writing group. Another result suggested by the rankings is that it is simply not possible to raise the level of poor arguers up to the level of the best arguers by influencing them to use better techniques. It would be interesting to conduct experiments using a within subjects approach, to see how consistently the same subject might produce higher ranked arguments with LEGALESE than with pen and paper.

It is possible to infer that use of LEGALESE does influence subjects to use the LEGALESE approach, and that the approach is more likely to lead to better arguments. There is also an indication that the LEGALESE subjects who performed the worst were subjects who simply did not grasp how to make use of the tool. This result suggests that there are problems with the interface itself, since it had been expected that all subjects would understand the approach of the tool after a brief introduction.

The fact that the LEGALESE subjects had a clear tendency to use complete sentences in the first phase reflects both the difference in the environment and an unfamiliarity with the LEGALESE environment. The writing group was performing a familiar act by constructing pen and paper outlines, and the act was one which is associated with the use of sentence fragments. On the other hand, the LEGALESE group was working with an unfamiliar tool. The computer setting itself may have suggested the use of complete sentences, since only

complete sentences appeared in the example and the setting could have struck subjects as calling for more formal structures.

Even more suggestive of using complete sentences is the fact that the LEGALESE tool does not provide much context for the entry of new argument pieces. When a user is able to edit a given piece of an argument the user sees little else to place the piece in its overall context. The writing of complete sentences, then, reflects an effort to make each piece stand on its own. In effect, complete sentences are an effort to supply context.

The fact that complete sentences were used suggests that users were more focused on each argument piece than was the case with the writing subjects. Thus, it presumably takes longer to compose complete sentences than sentence fragments.

These results indicate that the tool itself needs some improvement in two areas. First, it would be useful to establish more immediate context when users are editing individual pieces of argument structures. Second, through devices such as providing additional context it might be possible to make the ideas behind the tool's approach more readily understood by all users.

Conclusion

The work described in this paper was aimed at developing empirical evidence bearing on the question of whether human performance at high level cognitive tasks can be improved through the use of specially designed computer interfaces. While the evidence gathered is not entirely conclusive, the results discussed here do suggest that the LEGALESE tool has successfully improved people's performance at the task of constructing legal arguments.

References

- [1] Bush, V., As we may think, *Atlantic Monthly*, pp. 101-108, July, 1945.
- [2] Conklin, J., Hypertext: An introduction and survey, *Computer*, pp. 17-41, September, 1987.
- [3] Conklin, J. and Begeman, M.L., gIBIS: A hypertext tool for team design deliberation, *Hypertext'87 Papers*, pp. 247-251, Chapel Hill, North Carolina, 1987.
- [4] Engelbart, D.C., Augmenting human intellect: A conceptual framework, Summary Report Contract AF 49638 1024 SRI Project 3578, Stanford Research Institute, Menlo Park, California, 1962.
- [5] Hair, D.C., LEGALESE: A legal argumentation tool, to appear in *SIGCHI Bulletin*.
- [6] Lewis, C., Using the "thinking-aloud" method in cognitive interface design, IBM Research Report RC 9265 (#40713), 1982.
- [7] Marshall, C.C., Representing the structure of a legal argument, *Proceedings of the Second International Conference on Artificial Intelligence and Law*, pp. 121-127, 1989.

[8] McCall, R., PHI: An argumentative design method, to appear in *Design Methods and Theories*.

[9] Smolensky, P., Fox, B., King, R., and Lewis, C., Computer-aided reasoned discourse, or, how to argue with a computer, *Cognitive Science and Its Applications for Human-Computer Interaction*, Guindon, R. (ed.), pp. 109-162, Lawrence Erlbaum, 1988.

The following materials have been put together for you by your law clerk in connection with your representation of Mr. Sam Sincere. These materials represent the facts that the clerk thinks are relevant, along with citations to all of the statutes and precedents the clerk thinks may have bearing on this case.

FACT SITUATION:

Sam Sincere is a Viet Nam war veteran. As a result of stepping on a land mine while on night patrol in Viet Nam he lost both of his arms and both of his legs. His only means of getting around by himself is in a motorized wheelchair that he maneuvers with a mouth control.

On Veterans Day Mr. Sincere was asked to attend a ceremony in honor of Viet Nam veterans that was held in John Wayne State Park. That park is located in the city limits of Big City which is in the state of Colorado., As it happens, Big City is on the state line between Colorado and Wyoming, and Mr. Sincere is a resident of Wyoming.

As Mr. Sincere was traveling through the park in his wheelchair he was stopped by Officer Petty of the Big City police department. Officer Petty gave Mr. Sincere a citation for breaking the Big City ordinance forbidding the use of motor vehicles in a park. Mr. Sincere was unaware of the existence of the ordinance.

DESCRIPTION OF STATUTES AND CASES:

Officer Petty cited Mr. Sincere under section 5.01 of the Big City ordinances which reads as follows:

It shall be unlawful for anyone to operate a motor vehicle in any public park located within the city limits of Big City.

In the case of Big City v. Smith, the Colorado Supreme Court held that section 5.01 was a valid city ordinance and that it was properly applied against a one-legged man who drove a specially designed car into a park.

In the case of Big City v. Biker, the Colorado Supreme Court held that section 5.01 was properly applied against a woman going through a park on a motorcycle.

In the case of Big City v. Golfer, the Colorado Supreme Court held that section 5.01 was properly applied against a man who drove a golf cart into a park from an adjoining public golf course while looking for a golf ball.

In the case of Garden City v. Young, the Colorado Supreme Court held that an ordinance of Garden City, similar to the Big City ordinance, could not be applied against a person who was operating a soap box derby vehicle in a park.

Please use these materials, along with any other actual case law you may be aware of, to construct an argument in defense of Mr. Sincere.

Diagram 1. Hypothetical Legal Situation

Rank	Lawyer or Student	LEGALESE or Writing group	Subject number	Not Counted (Mann-Whitney)
(best)				
1.	student	LEGALESE	19	
2.	student	LEGALESE	2	
3.	lawyer	LEGALESE	12	
4.	lawyer	LEGALESE	15	
5.	student	LEGALESE	13	
6.	student	writing	18	
7.	student	LEGALESE	1	
8.	student	writing	16	
9.	student	writing	20	
10.	student	writing	3	
11.	student	writing	17	
12.	lawyer	writing	4	
13.	lawyer	LEGALESE	10	
14.	student	writing	22	
15.	student	LEGALESE	21	XX
16.	student	LEGALESE	23	
17.	student	writing	9	
18.	lawyer	LEGALESE	14	XX
19.	student	LEGALESE	8	XX
(worst)				

Table 1. Ranking Results.