The role of compression and refinement in visualization tools for crime analysts

Susan W. van den Braak, Herre van Oostendorp, Gerard A.W. Vreeswijk,
Department of Information and Computing Sciences, Utrecht University, The Netherlands
Email: susanb@cs.uu.nl, herre@cs.uu.nl, gv@cs.uu.nl
Henry Prakken, Department of Information and Computing Sciences, Utrecht University
& Faculty of Law, University of Groningen, The Netherlands, henry@cs.uu.nl

Abstract: This paper presents the results of a study that tested the effect of compression and refinement, as implemented in the *AVERs* graph visualization software for crime analysts, on the quality of the users' analysis of a simple crime case and their understanding of this case. In this study professional crime analysts and students who used these methods outperformed users that were only allowed to use conventional methods to handle large graphs.

Introduction

Graph visualization tools have been shown promising for learning and for collaborating on the construction and evaluation of arguments (Kirschner, Buckingham Shum, & Carr, 2003; van den Braak, van Oostendorp, Prakken, & Vreeswijk, 2006). In this spirit, a tool for the graphical visualization of stories and evidence named *AVERs* (Argument Visualization for Evidential Reasoning based on stories) has been developed (Bex et al., 2007). This tools aims at crime analysts who may use it to construct stories about what happened by linking events into causal chains and to connect the available evidence with these stories through arguments.

Problem statement

When large arguments are produced graph visualizations are often hard to read. As soon as the size of the graphs or the link density increases, such graphs become increasingly more complex and harder to understand. Argument visualization software should therefore offer methods that allow users to display their graphs in a readable way. Specialized software for crime analysts should additionally offer methods to elaborate on graphs and make underlying reasons explicit. It should be easy to unfold all hidden information about a certain node if desired, while still being able to oversee the larger picture. This is necessary due to the interactive nature of the analysis process in which an enormous amount of information needs to be analyzed and the constructed graphs are continuously being expanded and refined, while the reasons why certain links are added are often left implicit. We suggest a combination of compression and refinement as a viable solution, since it improves the readability of the graphs, while it also allows for elaboration. Compression is based on the idea that sometimes lines of reasoning are compressed into a rule. Take for example, the two-step argument for believing witnesses: "witnesses that speak the truth should be believed" and "witnesses normally speak the truth". This may then be compressed into the rule "witnesses should be believed". If a rule is to be attacked, it has to be restated in an uncompressed form. It is then easy to see that this rule can be attacked by arguing that witnesses who have a reason to lie do not speak the truth and therefore should not be believed (Loui & Norman, 1995). While decompression is used to add reasons for questionable links, refinement allows for the addition of more detail to earlier established links. We suggest that useful software tools for crime analysts should provide a combination of these two methods, since they correspond to tasks that are important during crime analysis.

Method

The goal of the study was to determine whether tools that contain refinement and compression methods support their users better than tools that provide conventional methods to handle large graphs. The study was conducted during a three-hour session at the Dutch police academy in Zutphen; 5 students of crime analysis and 11 analysts working in different districts in the Netherlands participated. Subjects were assigned to the conditions randomly. The treatment group was allowed to use a system which contained refinement and compression methods in order to analyze a simple case. The control group analyzed the same case by using a basic system which contained a simpler method for collapsibility that allowed them to collapse or expand all nodes of a certain type all at once, but in contrast to the treatment condition it was impossible to expand or collapse individual nodes. We predicted that the subjects in the treatment group would perform better, regarding the quality of their analysis and their understanding of the case, than the subjects in the control group.

A pre-test was conducted to help to account for biases between the treatment and the control group, after which the participants were asked to complete some instructional exercises to familiarize them with the system. Subsequently, in the actual test they had to analyze a simplified murder case using the system. After handing in their answers, a post-test was administered which consisted of true or false statements to test their understanding of the case. Finally, they were asked to fill in a usability questionnaire of 5-point Likert scale statements to measure the user-friendliness of the system as a whole and ease of use of specific features in it.

Results

Pre-test scores revealed that there were no significant pre-existing differences between groups. The graphs that were produced by the treatment group were more complete, better structured and more sound than the graphs produced by the control group, while the subjects in the control group (M=3972 seconds with SD = 878) used more time than the subjects in the treatment group (M=3286 seconds with SD=725). T-tests showed that the difference in soundness was significant (p=.04) and the difference in time nearly significant (p=.06), but the other differences were not (p=.40 for completeness and p=.24 for structure), although they were in the expected direction. In total, the graphs of the treatment group were better than the graphs of the control group (M=1.22 with SD=0.40 and M=0.89 with SD=0.41 respectively). This difference was marginally significant (p=.07). On the whole the data suggest that the treatment group produces higher quality analyses than the control group. The subjects in the treatment group also performed better on the true or false statements of the post-test than the control group (M=14.88 with SD=1.55 and M = 14.14 with SD=1.07 respectively). However, the difference found was not significant (p=.16). It should be noted that in the treatment group 4 subjects were able to answer all questions correctly, in the control group none of the subjects was able to do so.

The usability questionnaire revealed a mean rating of 2.80 (SD=0.60) on a 5-point scale for user-friendliness and a mean rating of 3.31 (SD=0.92) on the same scale for ease of use. More specifically, the subjects in the treatment group (M=3.63 with SD=0.92) found the collapsibility feature easier to use than the subjects in the control group (M=3.38 with SD=1.06), but this difference is not significant (p=.31). These results indicate that the user-friendliness of the system needs improvement as a satisfactory score should be at least higher than 3 on a scale from 1 to 5.

Conclusion and discussion

The study showed that crime analysts who are allowed to use methods to refine or compress links produce higher quality analyses and understand the case better than analysts who are provided with simpler methods to handle large graphs. All differences found between groups were in the expected direction, but only the difference in soundness was statistically significant, while the differences in the overall quality of the produced graph and the time taken to complete the task were nearly significant and showed a trend in the predicted direction. On the whole the analyses presented in this paper indicate that the selected methods increase performance and they have shown the importance of suitable ways to handle large and complex graphs. The usability measures revealed that the ease of use of the features in *AVERs* is satisfactory but that the user-friendliness of the system as a whole needs improvement. In particular, with respect to user-friendliness the inability to undo actions was pointed out as a drawback. While devising future versions of the system we will pay extra attention to this area.

Although the results presented in this paper are promising and in the predicted direction, the effects were not strong. Two reasons for this may be identified. Due to time constraints the case that had to be analyzed was rather small and simple. Arguably, in larger cases the differences between conditions might be even more apparent. Additionally, we expect that repetition of this study with a larger number of subjects will yield more significant results. Nonetheless this study provides preliminary support for the claim that compression and refinement indeed support users better than conventional methods to produce readable graphs, while they also satisfy the specific needs of crime analysts.

References

Bex, F., Braak, S. van den, Oostendorp, H. van, Prakken, H., Verheij, B., & Vreeswijk, G. (2007). Sense-making software for crime investigation: How to combine stories and arguments? *Law, Probability and Risk, 6*(1-4), 145-168.

Braak, S. W. van den, Oostendorp, H. van, Prakken, H., & Vreeswijk, G. A. (2006). A critical review of argument visualization tools: Do users become better reasoners? In F. Grasso, R. Kibble, & C. A. Reed (Ed.), *Workshop Notes of the ECAI-2006 Workshop on Computational Models of Natural Argument* (pp. 67-75). Riva del Garda, Italy.

Kirschner, P., Buckingham Shum, S. J., & Carr, C. S. (2003). *Visualizing argumentation: Software tools for collaborative and educational sense-making.* London, UK: Springer.

Loui, R. P., & Norman, J. (1995). Rationales and argument moves. *Artificial Intelligence and Law, 3,* 159-189.

Acknowledgments

This research was supported by the Netherlands Organisation for Scientific Research (NWO) under project number 634.000.429.