Sense-making software for crime investigation: how to combine stories and arguments?[1]

Floris J. Bex [a] Susan W. van den Braak [b] Herre van Oostendorp [b]
Henry Prakken [ba] Bart Verheij [c] Gerard A.W. Vreeswijk [b]

[a] *Centre for Law & ICT, University of Groningen*
[b] *Department of Information and Computing Sciences, Utrecht University*
[c] *Department of Artificial Intelligence, University of Groningen*

Crime investigation is a difficult and laborious process, especially in large cases that involve a mass of unstructured evidence of which crime investigators have to make sense. Because of the difficulty of this task and the high costs of mistakes, it is worth investigating how crime investigators could benefit from support tools. Current professional state-of-the-art software for managing and visualizing evidence, such as Analyst's Notebook[2], has the limitation that it does not allow for expressing the reasons why certain pieces of evidence support or attack a certain hypothesis. Therefore, in this paper a reasoning model for proof-of-concept software is proposed that offers crime investigators the ability to visualize their reasoning about a case. Subsequently, an implementation of this model in the *AVERs* (Argument Visualization for Evidential Reasoning based on stories) prototype system is presented. Crime investigators can use this sense-making tool to construct possible stories about what happened and link these stories with the available evidence through arguments. Arguably, such software has good chances of being useful while based on sound theories of the reasoning involved in crime investigations.

To be usable in practice and to improve the quality of crime investigations, our software design should be based on a natural and rationally well-founded theory of reasoning about evidence. To this end, two well-known accounts of reasoning about evidence, namely the argumentative approach as advocated in the New Evidence Theory (NET) [1] and the story based approach of Anchored Narratives Theory (ANT) [2] will be combined. The former takes its inspiration in Wigmores charting method, in which arguments from evidence to conclusions can be visualized. The latter stresses the importance of constructing stories about what might have happened in a case and of "anchoring" them in commonsense generalizations. In our opinion, both approaches have their shortcomings: NET does not allow for the comparison of stories and ANT does not give an accurate account of how stories can be connected to evidence. Our aim in this paper is to formalize and combine the two theories; this solves some of the problems of the separate theories and clarifies the relations between stories and arguments. The model should incorporate two ways of reasoning with causal knowledge, that is, from cause to effect (causal generalizations, e.g. fire *causes* smoke) and from effect to cause (evidential generalizations, e.g. smoke *is evidence for* fire).

In our model, the different stories about what happened in a case are represented as networks of causal generalizations. Such a causal network can be used as a causal theory, in which observations are explained by hypothesized events or states through abductive inference to the best explanation (IBE). These observations are inferred from the evidence using evidential generalizations. Thus, the evidence is linked to the story using arguments. The combined formal theory contains an *abductive framework* and an *evidential framework*. This combination is motivated by the way in which crime analysts in practice often combine time lines with sources of evidence.

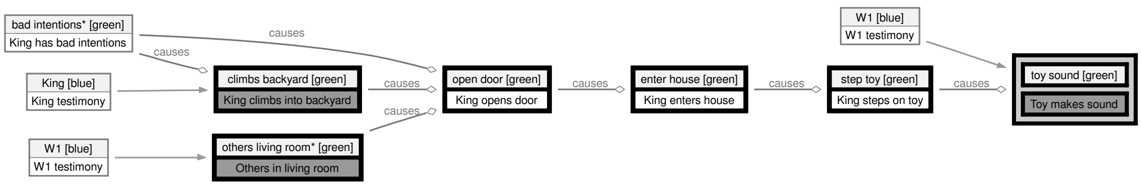
The *abductive framework*, which is a variant of logical AI models of abduction, is a tuple $A_C = (G_C, O, F, X)$. G_C is a set of *causal generalizations* of the form g: p_1 & $\ldots$ & $p_n \Rightarrow_C q$; the set O

[1] The full version of this paper will appear in *Law, Probability and Risk* and is available from `http://www.cs.uu.nl/research/projects/evidence/publications.html`. This research was supported by the Netherlands Organisation for Scientific Research (NWO) under project number 634.000.429.

[2] `http://www.i2.co.uk/Products/Analysts_Notebook/default.asp`

contains the *observations* which are the events in the story that are supported by evidence; the set $F \subseteq O$ contains the *explananda*, which are the observations that need to be explained and finally, the set X contains the *explanantia*, which are the propositions that can be part of a hypothesis. Now, an *explanation* in terms of A_C is a consistent set $H \subseteq X$ that explains all propositions in F. If there are more explanations that explain F, they can be compared with each other by looking at how many of the observations *not* in F (*additional evidence*) are explained or contradicted by the different explanations. The *evidential framework*, which is a variant of logical AI models of defeasible argumentation, is a tuple $A_E = (G_E, E)$, where G_E is the set of *evidential generalizations* of the form g: p_1 & ... & $p_n \Rightarrow_E q$ and E is the set containing the evidence. The arguments that can be built with the elements from G_E and E influence the abductive explanations in two ways. Firstly, the observations in O have to be inferred from the evidence E using generalizations from G_E. Secondly, it is possible to influence the content of the set G_C by arguing about the validity of the causal generalizations.

The outlined model is implemented in the prototype system *AVERs*; here illustrated through the King case[3] in which King is accused of climbing into the backyard of a family to rob them. However, he gets caught entering their house, because he steps on a toy causing it to make a sound. Using this tool users can visualize stories and evidential arguments in graph-like structures. Stories can be constructed by linking claims about a case, represented as green boxes, through causal links, which are yellow with diamond-shaped arrowheads. Secondly, it allows stories to be connected with the available evidence. To do this evidence may be added by selecting text from source documents; such quotes are represented as blue boxes. These may then be linked to claims through evidential links that are represented as blue arrows; the system automatically adds these supported claims to the set of observations O. Moreover, the system provides support for evaluating and comparing stories based on the number of observations that are explained by a certain hypothesis. Given a set H of nodes marked by the user as hypotheses (denoted by an asterisk in the graph) and the set of causal generalizations G_C drawn by the user, *AVERs* determines which of the observations in O are explained by $H \cup G_C$. A good hypothesis should explain as many events and observations from O as possible but at least the nodes in F. The system provides visual feedback to the user by displaying the nodes that are part of the set O in a gray color and the nodes in F as encapsulated boxes. Additionally, when a certain hypothesis is selected by the user, the system will mark explained observations by a black box. In this way a user can easily determine the quality of his story and compare it to other alternatives.



We claim to have made two main contributions. Firstly, we have argued that in the context of sense-making systems for crime investigation there are reasons to combine two AI approaches, abductive IBE and defeasible argumentation, which are usually considered as irreconcilable alternatives. Secondly, we have described the current design of a visualization software tool in which causal networks can be linked to the available evidence with argumentation structures. At later stages of the project it will be empirically tested in user experiments to see whether using such a tool indeed has the benefits it is often claimed to have. Among other things, this should also bring clarity on whether a combination of story-based and argument-based reasoning is indeed natural to crime investigators.

References

[1] Terence Anderson, David A. Schum, and William Twining. *Analysis of Evidence*. Cambridge University Press, 2nd edition, 2005.

[2] Willem Albert Wagenaar, Hans F.M. Crombag, and Peter J. van Koppen. *Anchored Narratives: Psychology of Proof in Criminal Law*. St Martin's Press / Prentice-Hall, New York, NY, 1993.

[3] Adapted from the original paper. The software design makes extensive use of colors, which cannot be shown here. Therefore, color indications are provided between brackets. The paper available online contains color figures.