

A Proposal for Evidential Reasoning about Motives

Floris Bex
University of Groningen
Faculty of Law
The Netherlands
f.j.bex@rug.nl

Katie Atkinson
University of Liverpool
Department of Computer Science
Liverpool L69 3BX, UK
katie@liverpool.ac.uk

ABSTRACT

Motives play an important role at every stage of a criminal investigation. In this research abstract we provide an overview of an account of motivations based on a general approach to practical reasoning.

1. INTRODUCTION

Motives play an important role at every stage of a criminal case. They can be used to search for an explanation of the crime (why was this person killed?), to identify a suspect (who would have the motive to kill this person) or to persuade a jury of a suspect's guilt (this motive explains why the suspect committed the crime). In this research abstract we discuss a framework for the analysis of reasoning with such motives and their underlying values, concentrating on the use of motives to provide plausibility to a story intended to persuade a jury of a person's guilt or innocence.

A persuasive story should be plausible in that it conforms to our beliefs about how things generally happen in the world around us. This plausibility is partly dependent on the plausibility of the (physical) causal relations between the events in the story. For example, a story in which the victim ends up in a coma because she was injected with insulin is only plausible if we believe that an overdose of insulin can cause a coma. The plausibility of a story is also dependent on how likely it is that the agents in the story would have made the alleged choice in the situation, and this in turn depends on the value preferences we believe an agent to have. For example, a story in which the suspect kills the victim in order to inherit the victim's money is only plausible if we believe that the suspect values money higher than the life of the victim. When determining the plausibility of a story we can thus differentiate between plausibility as regarded from the *physical* as well as the *intentional stance* [6].

Other work on evidential reasoning in AI and Law, most notably by Bex and colleagues [5] and Thagard [7], mainly focuses on reasoning from the physical stance. While both approaches allow for the inclusion of links that denote some

sort of "motivational causation" (i.e. motive *causes* action), the reasoning about why a certain choice was made by the agents in a case remains implicit. Similarly, Walton and Schafer [8] establish the existence of a motive, but lack the machinery to explain why an actor chose to act on this motive in the particular situation.

Our account of reasoning with motives and values is based on a general approach to practical reasoning that was previously introduced in [3]. Here we discuss how explanations of what happened in a criminal case can be inferred through abductive reasoning, and how the possible motives of the agents can influence the choice between these explanations.

2. MOTIVES IN EVIDENTIAL REASONING

When discussing reasoning about motives, it is useful to distinguish the separate concepts of *value*, *motive* and *goal*. Values, as defined by [2], can be seen as abstract principles which an agent or a group of agents hold. Examples of values are "wealth", "love" and "honesty". Agents can be expected to, actively or passively, promote values which they find important. Our concept of motive is roughly the same as this concept of value, in that a motive is an abstract good which an agent may or may not want to promote. Such a motive can cause a person to form any number of goals in order to promote the principle for which the motive stands. For example, if a husband is motivated by the will to be independent, he can plan to murder his wife. However, he can also form the less drastic plan to simply divorce his wife. So the same motive can cause a person to form different goals. This current notion of motive and goal is somewhat different from in [8], where an agent's motive is essentially equated with an immediate goal of the agent. In our opinion, our notion of motive better captures the basic intuition that the same motive can be satisfied in different ways.

In all stages of criminal investigation it is important to consider possible alternatives so as to avoid the well-known problem of confirmation bias or tunnel vision. One way of providing such an alternative is to tell a totally new story (with possibly a different suspect) that explains the evidence at least as well as the current story. Another way of providing an alternative is to argue that this suspect was not the perpetrator of the crime. Instead of providing a totally new explanation for the evidence, it is simply argued that the current suspect could not or would not commit such a crime, thus providing what we call a "suspect-specific" alternative; for example, it can be argued that it was not physically possible for the suspect to have committed the crime because he was somewhere else at the time, thus providing

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICAIL-2009 Barcelona, Spain

Copyright 2009 ACM 1-60558-597-0/09/0006 ...\$5.00.

him with an alibi. It can also be argued that the suspect is not the type of person who would commit such a crime by showing that his motivational preferences are different from those assumed; for example, one could argue that the husband values love for his wife higher than money so he would never kill her, even though he would benefit financially from her death. Finally, it can be argued that given the suspect's values, he could have satisfied his motives by acting differently; for example, arguing that it would have been easier for the suspect to divorce his wife and take half her fortune.

These last two ways of providing a suspect-specific alternative are examples of reasoning about the suspect's motives and the suspect's willingness to act on such motives. This shows that in any (formal) reasoning framework which is used in evidential reasoning we should be able to rationalise why the suspect in a case chose to act on his motives in that particular situation. A notion related to this is that of *character*, which can be modelled as an ordering of a person's motives. For example, we can say that X is the kind of person who prefers independence over honesty. If we have evidence that X is such a person - for example, that he has lied on previous occasions - we can provide an argument for this particular preference. As Walton and Schafer [8] note, character evidence is often inadmissible at trial. However, it may be admissible for the purpose of proving someone's motive, which is exactly the purpose we wish to use it for.

3. PRACTICAL REASONING WITH MOTIVES

In [3], we attempted to analyse motives in terms of an agent choosing to act so as to promote some value by realising some goal, an approach to practical reasoning previously developed in [1]. The approach is based on the use of an argumentation scheme and critical questions which allow justifications for action to be presumptively proposed and critiqued. The scheme can be used by an agent to justify an action in particular circumstances in terms of his values. For a given instantiation of this scheme a number of critical questions could be posed to challenge the various elements of the scheme and so dispute the presumptive conclusion.

In order to provide a formal basis for this argumentation scheme, the problem scenario is modelled as an Action-based Alternating Transition System (AATS) [9]. Essentially, an AATS consists of a set of states and transitions between them representing the possible joint actions of the agents in a given state. These transitions are labelled with the values (motives) promoted by moving from the source state to the target state, and the notion of a goal is identified with the new state. The argumentation scheme and all of its associated critical questions are given formal definitions in terms of an AATS in [1]. Arguments following this scheme are now based on a particular path from the current state to a new state in the AATS. As an agent can only perform one action in the given circumstances, arguments for different actions attack each other. The critical questions also point to possible counter-arguments. Once a set of arguments and the attacks between them has been generated on the basis of a specific AATS, the status of the arguments can be evaluated. To do this, the arguments are formed into a Value-Based Argumentation Framework (VAF) [2].

In [3], we extended the above non-abductive view on practical reasoning with a separate argument scheme for abductive practical reasoning to allow for search and investigation cases. This abductive scheme makes it possible to infer an

explanation for a set of circumstances in terms of a motivated action, and explanations instantiating the scheme can again be critiqued. The abductive scheme can be combined with the normal scheme, which enables us to reason both predictively and explanatorily about motivated actions.

One of the key elements our approach brings is the explicit distinction between arguments that can be made to reason about physical causal relations and arguments that can be made to reason about motivations and their priorities. The framework described here allows us to predict what possible actions an agent would take given his values and an ordering on these values. This kind of reasoning is often employed to argue that, given the circumstances, the suspect would have acted in a particular way to fulfill his motives. Since crimes are "by definition deviant behaviour, what people do not normally do" [8], the key to the plausibility of this argument is explaining why the accused acted in the aberrant fashion. While a typical agent would not have chosen the action using a "normal" value order, there is some value order on which it would be chosen: that the agent under suspicion in fact had this unusual ranking of values itself needs to be explained in order to make the agent's choice of this action plausible. Consequently, when considered from the intentional stance, there are three elements to a plausible story: 1) a motive for the action, 2) an explanation of why the agent had this motive, and 3) an explanation of why the agent's value order was such as to make this motive of sufficient importance. The approach has been applied and an actual case modelled in [4]. The case modelled is that used by [7], allowing specific comparison with Thagard's approach.

4. REFERENCES

- [1] K. Atkinson and T. J. M. Bench-Capon. Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artificial Intelligence*, 171(10–15):855–874, 2007.
- [2] T. J. M. Bench-Capon. Persuasion in practical argument using value based argumentation frameworks. *J. of Logic and Computation*, 13(3):429–48, 2003.
- [3] F. J. Bex, T. J. M. Bench-Capon, and K. Atkinson. Did he jump or was he pushed? Abductive practical reasoning. *Accepted for publication in Artificial Intelligence and Law*, 2009. Available online.
- [4] F. J. Bex, T. J. M. Bench-Capon, and K. Atkinson. Evidential reasoning about motives: A case study. Technical Report ULCS-09-013, Dept of Computer Science, University of Liverpool, UK, 2009.
- [5] F. J. Bex, H. Prakken, and B. Verheij. Formalising argumentative story-based analysis of evidence. In *Proceedings of ICAIL 2007*, pages 1–10, 2007.
- [6] D. Dennett. Intentional systems. In *Brainstorms*, pages 3–22. Harvester Press, 1979.
- [7] P. Thagard. Causal inference in legal decision making: Explanatory coherence vs. bayesian networks. *Applied Artificial Intelligence*, 18:231–241, 2004.
- [8] D. Walton and B. Schafer. Arthur, George and the mystery of the missing motive: Towards a theory of evidentiary reasoning about motives. *International Commentary on Evidence*, 4(2):1–47, 2006.
- [9] M. Wooldridge and W. van der Hoek. On obligations and normative ability: Towards a logical analysis of the social contract. *J. of Applied Logic*, 3:396–420, 2005.