

Arthur, George and the Mystery of the Missing Motive: Towards a Theory of Evidentiary Reasoning about Motives *International Commentary on Evidence*, 2006, 4(2), 1-45.

Douglas Walton, University of Windsor & Burkhard Schafer University of Edinburgh

## Abstract

Reasoning about motives is a prominent part of the investigative process in Julian Barnes' *Arthur and George* and in Conan Doyle's novels themselves. In *Arthur & George*, it is the lack of motive that originally convinces Arthur of George's innocence, a conclusion that a subsequent evaluation of George's character supports. In many of the Sherlock Holmes novels, it is the discovery of a hidden motive that puts Holmes on the right track. On closer inspection though, the evidentiary value and logical structure of arguments from motive turn out to be problematic. In this paper, we use concepts from argumentation theory and computational models of agent behavior from artificial intelligence research to analyze the structure of motive-based reasoning in fictional and factual crime investigations and trials. This analysis is used to develop a theory that (a) accounts for the rationality of motive-based reasoning, (b) helps to distinguish plausible and implausible arguments from motive, and (c) to distinguish arguments based on motive from arguments based on intention and character, as well as from other, legally more problematic forms of evidentiary reasoning.

**KEYWORDS:** motives, argumentation, Conan Doyle

### 1. Introduction: Reasoning about Motives in Fact and Fiction

#### 1.1 Sherlock Holmes on Motives

Conan Doyle's Sherlock Holmes stories are often credited with bringing the investigative trinity of "means, motive and opportunity" into the public mainstream (Zizek 1990). Without any doubt, reasoning about motives forms an integral part of Holmes' crime solving activities. "There is no crime to detect, or, at most, some bungling villainy with a motive so transparent that even a Scotland Yard official can see through it" he says in *A Study in Scarlet*. Simple crimes are simple precisely because the motive is so obvious that even the police can detect it, and through it solve the crime. The hallmark for the truly great investigator by contrast is his ability to find the non-obvious motive and in this way to solve even the most difficult of cases. An even better idea of the role and function of motives in Holmes' reasoning can be found in *A Case of Identity*:

[The cases given to me] are important, you understand, without being interesting. Indeed, I have found that it is usually in unimportant matters that there is a field for the observation, and for the quick analysis of cause and effect which gives the charm to an investigation. The larger crimes are apt to be the simpler, for the bigger the crime the more obvious, as a rule, is the motive. In these cases, save for one rather intricate matter which has been referred to me from Marseilles, there is nothing which presents any features of interest.

This quote gives us, so to speak, the basic ontology of Holmes' world. The "trifles" or traces, the footprints, cigarette ash and wounds on bodies that are the starting point of his investigative activity are of interest to him only because they have been *caused* by rational human agents in their *intentional* pursuit of their goals. Holmes is not a forensic scientist, he is a detective. For the forensic scientist, it is sufficient to establish the immediate or physical cause of a trace.

The presence of ash can be explained by someone smoking. For the detective however, this is only part of the picture. A fuller explanation for him would be: The perpetrator was waiting for his victim. To kill the time and fight his nervousness (indicating premeditation) he smoked a cigarette, which left the traces on the crime scene. This inference is with necessity more speculative than that of the forensic scientist, but also richer in content. Again we can quote Holmes, this time in conversation with Watson:

"Yes," said I. "You have not made it clear what was Colonel Moran's motive in murdering the Honourable Ronald Adair? Ah! my dear Watson, there we come into those realms of conjecture,

where the most logical mind may be at fault. Each may form his own hypothesis upon the present evidence, and yours is as likely to be correct as mine." "You have formed one, then?" "I think that it is not difficult to explain the facts. It came out in evidence that Colonel Moran and young Adair had, between them, won a considerable amount of money. Now, undoubtedly he played foul--of that I have long been aware. I believe that on the day of the murder Adair had discovered that Moran was cheating. Very likely he had spoken to him privately, and had threatened to expose him unless he voluntarily resigned his membership of the club, and promised not to play cards again. It is unlikely that a youngster like Adair would at once make a hideous scandal by exposing a well-known man so much older than himself. Probably he acted as I suggest. The exclusion from his clubs would mean ruin to Moran, who lived by his ill-gotten card-gains. He therefore murdered Adair, who at the time was endeavouring to work out how much money he should himself return, since he could not profit by his partner's foul play. He locked the door lest the ladies should surprise him and insist upon knowing what he was doing with these names and coins. Will it pass?" "I have no doubt that you have hit upon the truth."  
(*The Adventure of the Empty House*)

The semiotic analysis of Holmes' reasoning that Sebeok and Eco (1988) made popular falls short in a crucial aspect: traces are not just signs of the action that brought them about, they are also indicative of the intention or motive behind that action. In *A Study in Scarlet*, this quote from a paper written by Holmes is given:

"From a drop of water," said the writer, "a logician could infer the possibility of an Atlantic or a Niagara without having seen or heard of one or the other". So all life is a great chain, the nature of which is known whenever we are shown a single link of it. Like all other arts, the Science of Deduction and Analysis is one which can only be acquired by long and patient study nor is life long enough to allow any mortal to attain the highest possible perfection in it. Before turning to those moral and mental aspects of the matter which present the greatest difficulties, let the enquirer begin by mastering more elementary problems. Let him, on meeting a fellow mortal, learn at a glance to distinguish the history of the man, and the trade or profession to which he belongs.

Physical and mental aspects of action are inseparable in this model, the chain of reasoning leads inevitably from one to the other.

## 1.2 From Holmes to "Arthur": On Motives and Conspiracies

How does Julian Barnes "Arthur" measure up to Sherlock Holmes? The answer is somewhat mixed. Barnes' description of the first meeting between Arthur and George is very interesting for our purpose. At that point in the novel, Arthur has studied the dossier George compiled. On the basis of the dossier, he has formed a firm conviction of George's innocence, a conclusion further strengthened through his meeting with George.

One obvious way to account for the rationality of Arthur's conviction is as a Holmesian abductive inference from the established evidence, the letters and the mutilated pony, to the *intentional, goal-directed* action of a human being. George could not have had any conceivable motive for carrying out actions that produced the established evidence. Granted, an insane person may have mutilated the pony for whatever dark desires drove him. But there was no evidence to support the claim that George was of that disposition, and his functioning in a demanding job would indicate otherwise, as did Arthur's direct assessment of George's character during their first meeting. Even if George had some inexplicable motive for mutilating the ponies, it would not explain (and, indeed, would be inconsistent with, the production of the accusatory letters, or the established evidence of a long lasting campaign of harassment. Following Holmes's maxim that "that when you have excluded the impossible, whatever remains, however improbable, must be the truth." (*The Adventure of the Beryl Coronet*), this absence of any plausible motive that would explain all the available evidence would indeed be enough to eliminate George from the list of suspects.

We can see here how motive evidence is, for Holmes, of much the same significance as evidence of "means" and "opportunity": To establish that I had means and opportunity to commit a crime adds little

weight to the prosecution case. At any given time, this will be true for many people. But the absence of opportunity (alibi) or means (no in possession of the right type of weapon) rules me out with logical necessity. In Holmes' world (as we will see, in marked distinction from the way that the law approaches the issue of motive) there is no action without motive, and absence of motive is close to a logical proof of innocence. In argumentation terms, absence of motive can therefore function as a defeater to even very compelling forensic evidence.

However, Arthur is not content with relying on the lack of motive to eliminate George from the class of possible perpetrators of the crime at issue. Not only did George lack a motive, others had an abundance of motives that could explain the case. Chief among them are the scorned and dismissed maid (motivated by her desire of revenge), Sergeant Upton (motivated by the most glaring racism), and his superior, Captain Hanson (motivated by the desire to maintain the reputation of the force and to protect "his people" against the outsider). Even so, Barnes' George dismisses Arthur's theory of a racially motivated conspiracy against George out of hand: "But I'm a lawyer, Sir Arthur. What evidence do I have that anyone has acted against me because of race prejudice?" In insisting on evidence to support the motive argument, George dons the gown of Sherlock Holmes. "The temptation to form premature theories upon insufficient data is the bane of our profession", Holmes said in the *Valley of Fear*, and "It is a capital mistake to theorize before one has data. Inevitably one begins to twist facts to suit theories, instead of theories to suit facts". Between them, Arthur and George may have made a passable Sherlock.

Arthur's premature assumption of a racially motivated conspiracy, even if correct in George's case, points to one of the problems with reasoning from motives: every good conspiracy theory employs the same mode of reasoning. We observe that certain people, or groups of people, benefit from a development. From this we infer that they had a motive. From this, it follows that they brought about the observed result through conspiring with each other. This is the most basic pattern of conspiracy theories. A convincing theory of reasoning with motives should also establish robust criteria to distinguish problematic conspiracy theories from appropriate reasoning about collective motives and benefits. In our approach, the concept of "critical question" is designed to take on that role. However, to fully develop such a set of critical questions requires a more substantive engagement with the way in which problematic conspiracy theories can be identified than can be given here, and must remain the topic of a future paper.

A very short indication of the type of reasoning that would have to be modelled should nonetheless be given to indicate broadly the type of approach intended. Popper (1966 pp.94-96) was a staunch critic of conspiracy theories as methods of explanation, and for a long time the prevailing opinion in epistemology and theory of social sciences seemingly accepted his criticism as valid. Popper characterized the conspiracy theory of Society as the view that:

[...] an explanation of a social phenomenon consists in the discovery of the men or groups who are interested in the occurrence of this phenomenon (sometimes it is a hidden interest that first has to be revealed) and who have planned and conspired to bring it about (Popper 1966 p. 94).

"Arthur's" position is clearly a conspiracy theory. Undoubtedly it has its dangers as a pattern of explanation. On the other hand, it seems equally obvious that sometimes, there are conspiracies – Pigden in his reply to Popper offers an impressive list of social phenomena that historians regard as the results of conspiracies, from Elizabethan times to Watergate, (Pigden 1995). Criminal law in many countries also reflects the belief that conspiracies can exist. Criminal law governing a range of phenomena that include membership in criminal organizations and illegal trade cartels requires that prosecutors show that several people had converging motives, which caused them to bring about a certain state of affairs (Hughes and Sanderson 2004). In international human rights law and international criminal law in particular, this type of analysis plays an important role (Farer 2000). In international human rights law, the state itself is called to account, and the prosecution has to show as a result a systematic and coordinated activity of several of its agents. In international criminal law too, more complex patterns are the subject of investigation, not individual and isolated actions. International crimes require typically systemic, serious human rights violations which occur in a context of mass violence (such as a widespread or systematic attack directed against a civilian population or a manifest pattern of similar conduct. To show coordinated actions by several agents or "conspirators" is therefore frequently a prosecution requirement. It seems therefore that an adequate theory of evidentiary reasoning about

motives cannot take Popper's unforgiving stance, but rather should allow us to distinguish reliable and unreliable conspiracy theory arguments in a legal setting. Suitable modifications of Popper's theory have been proposed by Coady (2006) and Pigden (1995).

From the other end, Clarke (2006) has employed a perspective from cognitive science to develop a deeper analysis of the possible fallacies of conspiracy arguments. However, reasoning about the converging motives of several actors poses quite significant additional difficulties, both in law and logic. In particular, modeling that reasoning would require analysis of the concept of "having the same/ a converging motive as". Our analysis therefore concentrates on single actor cases. We do, however, try to avoid any commitment that would make an extension to multiple actor scenarios obviously impossible.

### 1.3 From Fiction to Fact: Motive Evidence in Forensic Argumentation and the Law

While Doyle learned to reason abductively from observed signs to physical causes during his studies under the Scottish surgeon Joseph Bell, he may well have picked up the importance of motives during his classical studies. Reasoning about motives has a long and distinguished history in forensic rhetoric, going back at least to Cicero:

There can be little foundation for a motive for crime unless suspicion is cast on the character of the defender so that it will not seem to be inconsistent with the act. For as there is no point in discrediting the character of a man where there was no motive for him to go wrong, so it is idle to allege a motive for a crime if his character is shown to be inclined to no line of conduct which is less than honorable. Therefore the prosecutor ought to discredit the life of the defender on the basis of his past acts and to point it out if he has previously been convicted of any crime equally serious. If this is impossible, he should prove that the defender has been under suspicion of similar crime before. . . . If it cannot be shown that the defender has ever been implicated in any fault, the argument will be brought in by which the jurors are urged to think that the long-standing reputation of the man has nothing to do with reality. For he has been concealing his true character before, and has only now been caught red-handed; therefore this act should not be judged in view of his past life, but his past life should be discredited by this act; that previously he had no power or motive to commit the crime. And if none of these arguments can be made, one should have recourse to the last possible argument, that it is no wonder that he has now for the first time committed a crime: for one who wishes to sin, must have some first offence (Inv. 2.32, 34).

While forensic rhetoric very much kept alive the tradition of reasoning about motives, which still features prominently in courses about the law of evidence, there is a comparative scarcity of literature both in doctrinal analysis of evidence law and in evidence scholarship. Cicero's quote partly explains why.

First, motive evidence falls between two main topics of jurisprudence and evidence scholarship – the question of intent or *mens rea* on the one hand, the issue of character evidence on the other. Writers in either field typically start their investigation by distinguishing their topic from the issue of motive, which as a result often escapes attention. Nor has the question of motive resulted in the creation of interesting doctrinal rules or problematic precedents worth academic attention. In Scotland, Doyle's home jurisdiction, even today the rules on admissibility of motive evidence fill hardly a page. The Crown need not normally establish motive – according to the institutional writer Hume (1844 p. 245) – but evidence by the Crown that a motive was present is admissible.<sup>1</sup> As an exception to the general rule, certain crimes, such as "racially aggravated assault", require that the aggravating factor is proved beyond reasonable doubt. In the case of *Slater v HM Advocate* (1928 JC 94), the fact that Slater was a dealer in jewels and therefore had a motive for killing the owner of a valuable brooch was deemed admissible and relevant. High profile cases of Holmes's time in which motive evidence featured prominently include the case of Madeleine Smith and the trial of Dr Pritchard, the notorious Edinburgh poisoner. Both were published in the *Notorious Scottish Trials* series (Farmer 2003) and Doyle would have been acutely aware of them. On the other hand, apparent absence of motive is an important argument in favour of the defence (Walker and Walker 1983 p. 98).

<sup>1</sup> 1 *Nelson v HM Advocate* 1994 SCCR 192, per Lord-Justice Hope at 202 E-F. *HM Advocate v Pritchard* (1865) 5 Irv88

With certain crimes, the absence of motive seems to be more damaging for the prosecution than in others. Financial dishonesty in particular will normally require that the prosecution shows that the accused was financially in distress (Walker and Walker p. 99). In all Scottish cases where the evidentiary rules regarding motive are discussed, it seems to be clearly a case not of admissibility, but of attribution of probative weight.

Second, Cicero makes his skepticism over the motive argument clear – it seems to be “largely” or “mainly” rhetorical, close to an ad hominem argument attacking the character of the accused. Efficient possibly, rational hardly. To some extent, as we have already seen, Cicero has a point. As an argument for the prosecution, motive evidence seems very weak indeed. Consider the following example. The prosecution argues: “Mr Grimesby Roylott was heavy in debt. He stood to inherit heavily after the death of Miss Julia and Helen Stoner. He therefore killed Miss Julia, and attempted to kill Helen.” Why is this argument strengthening the prosecution case? If we assumed that arguments, to be convincing, had to be deductively valid, we would try to reconstruct this argument in the tradition of Hempel and Oppenheimer, as application of a governing law to a particular case:

(Most) people who are in financial difficulties kill their daughters to inherit.  
Grimesby Roylott was in financial difficulties.  
He therefore killed his daughter.

This is obviously wrong. Only very few people kill their children, or step children, for financial gain. Almost all of them can sleep soundly in their beds, even if the financial burden on their parents is heavy. Indeed, crimes by definition are “deviant” behavior, what people normally do not do. Finding a governing law is therefore always going to be problematic. Nor is it an argument for empathy. The Crown is not arguing: “If you put yourself in the shoes of Mr Grimesby, you would have done the same”. As a prosecution strategy, this is hardly plausible. On the other hand, we have seen that when used by the defense, motive has an important role to play. Even in the prosecution case, we can distinguish between more or less convincing arguments by motive. Nor can their importance be doubted for the everyday practice of court decision making. The jury instruction from the OJ Simpson case strikes the balance well:

Motive is not an element of the crime charged and need not be shown. However, you may consider motive or lack of motive as a circumstance in this case. Presence of motive may tend to establish guilt. Absence of motive may tend to establish innocence. You will therefore give its presence or absence, as the case may be, the weight to which you find it to be entitled.

However, can we say more about this? What is the “weight to which it is entitled”? To answer these and similar questions, we will now develop the first step of a rational theory of motive evidence, by bringing together the tradition of forensic rhetoric with that of modern argumentation theory and the computational theory of rational agents. We hope to show how that theory leads to a rational reconstruction of the way in which Doyle uses motive evidence, gives indications how to analyze and weigh arguments by motive, and also helps us to distinguish arguments by motive from other types of evidentiary arguments. We will show how misuse of arguments concerning motive, such as the conspiracy theories discussed above, can be avoided with a system of “critical questions” that take account of George’s objection against Arthur’s premature theorizing.

Eventually, a theory of evidentiary reasoning with motives along the lines that we suggest should help to

- Differentiate more accurately between character and motive evidence
- Help to identify obvious mistakes in reasoning with motives and generally help to check for their correctness
- Help to compare different explanations using motive evidence in the same case, and help decide between them
- Ideally, help us to attribute a rational degree of evidentiary weight to a specific argument by motive

## 2. From Forensic Rhetoric to Rational Argumentation Theory

We present a logical model, based on current developments in artificial intelligence and argumentation that shows how lawyers, judges and fact finders rely on motive by identifying it as a species of argument. We show how this logical structure fits a model of reasoning currently attracting much attention in artificial intelligence and argumentation theory called practical reasoning. In the conception of practical reasoning most current in artificial intelligence and argumentation theory, an entity called a rational agent in multi-agent computing (Wooldridge, 2002) carries out actions based on goals, knowledge of its circumstances, and plans. We identify the premises and conclusion of motive-based reasoning, determine what types of argument it fits, and develop criteria that can be used to help judge how strong or weak it is as evidence in a given case. Tools of argument diagramming, argumentation schemes, and other argumentation-based methods, are useful for modeling many aspects of legal argumentation (Schum, 1994; Gordon, 1995; Prakken, 2001a; Lodder, 1999; Verheij, 2003; Reed and Norman, 2003; Walton, 2003; Prakken, Reed and Walton, 2003; Bex, Prakken, Reed and Walton, 2003; Verheij, 2005). Another useful tool is the technology of plan recognition systems developed in computing that uses a plan inference framework (Carberry, 1990) to infer goals from an agent's actions, including the agent's current and previous utterances in a dialogue, and common knowledge (scripts). This paper applies these tools to the problems of building a model that can be used to analyze and evaluate motive evidence.

Our analysis is based on a plan recognition model in which there can be reasoning forward, often called top-down reasoning, from a motive to an action, and reasoning backward, often called bottom-up reasoning, from an action to motive (Walton, 1990). Bottom-up plan recognition works by drawing inferences from an agent's observed actions to construct a plausible account of its expected goals and beliefs. In such a backward sequence, actions of an agent serve as evidence in a database, and the system uses the evidence to reason backwards to the postulation that the agent had goals and beliefs about its circumstances that acted as its reasons for the action being examined. Our analysis of motive evidence in law follows Wigmore's (1931, p. 146) insight that there is double inferential step, the first one from action to presumed motive, and the second one from motive to an ultimate *probandum*.

Our analysis defines a motive as an immediate goal of an agent's planning to which the agent is strongly attached. Again we follow Wigmore (1931, p.146), who defined a motive as a specific emotion, like being angry about a previous perceived harm suffered at the hands of a person that leads to an action, Walton and Schafer: Evidentiary Reasoning about Motives 9 like burning down that person's house. In the model we develop, motives are immediate goals that an agent is strongly committed to and adopts as a mainspring for its actions. We define a motive as a special type of goal that an agent is taken to have in mind that leads by a sequence of practical reasoning leading forward to an action carried out by that agent. This definition is a more detailed one than the general one given by Wigmore (1931, p. 146), who defined a motive as a specific emotion or passion likely to lead to the doing of a particular act.

Part of our definition is that although an agent's motive is internal, it can be reconstructed hypothetically by a second agent using backward practical reasoning to make a conjecture about the goal that is taken to be the basis of the first agent's observed action. One agent infers that another agent with whom it is engaged in a dialogue (Prakken, 2001, 2001a) has a particular motive, by using the evidence collected so far in the dialogue (what that other agent has said and is known to have done) and then, at a specific juncture in the dialogue, explains an action attributed to the first agent by inferring the existence of a motive. Such a sequence of reasoning, we show, is a species of abductive reasoning from a set of data comprised of the facts of case to a hypothesis whereby the first agent explains the presumed goals, actions and beliefs of the first agent.

### 3. Motive Evidence and Character Evidence

As noted above, there is a problematic overlap between motive evidence and character evidence. Evidence that a defendant was involved in prior crimes can be admissible if introduced for one purpose, but inadmissible if introduced for a different purpose. For example, consider the argument that the defendant was convicted of prior crimes, therefore he must have committed the crime he is currently charged with. When the proffering party offers evidence to support such an argument, that evidence is inadmissible. As noted above, offering evidence to show that the defendant has a bad character and therefore that he committed the offense at issue, is impermissible. But if the same kind of evidence were used to prove the existence of motive, it can be admissible. For as indicated by Park, Leonard and Goldberg (1998, p. 163), although motive is not an issue in a criminal prosecution, motive can be relevant if the ultimate issue is one of criminal intent or the identity of the perpetrator. They cite (p. 164) the

example of evidence of prior assaults, admitted in an assault case under the heading of motive. In such a case, the prior assault could be used as evidence that the defendant disliked the person he is charged with assaulting. The dislike would provide him with a motive, and would be admissible in that light to show that he was more likely to have committed the crime than someone else, who also had an opportunity to commit the crime, but had no motive.

Federal Rule 404 states that character evidence is not admissible to prove conduct, with a number of exceptions. One exception states that evidence of other crimes, although not admissible to prove the character of a person in order to show that he acted in conformity with his character, may be admissible for various other purposes. The section of rule 404 that states these other purposes is worth quoting.

It may, however, be admissible for other purposes, such as proof of motive, opportunity, intent, preparation, plan, knowledge, identity, or absence of mistake or accident, provided that upon request by the accused, the prosecution in a criminal case shall provide a reasonable notice in advance of trial, or during trial if the court excuses pretrial notice on good cause shown, of the general nature of any such evidence it intends to introduce at trial.

In law, the special problem is to show how the notion of motive is different from that of character. Park, Leonard and Goldberg (1998, p. 165) comment that evidence of other misconduct has been stretched, particularly in sexual misconduct cases, to the extent that there is a blurring of the line between evidence offered to show character and evidence offered to show motive. In general, it does not seem to be too easy to draw a firm analytical line between evidence as to character and evidence as to motive as distinct types of intermediate inferences that help establish the ultimate probandum. Part of the difference is that character tends to be thought of as the cause of habitual action, whereas motive often seems to play a role more like the mainspring of a particular action. Yet both character and motive represent common ways of explaining a given action, or trying to make sense of it, when the act is puzzling in some way, or when one is trying to determine something about the action.

Another initial basic problem, for both legal and cognitive science analysis of motive is the difficulty of defining the concept of motive is to define it in a way that shows how it differs from the related notion of intention. In artificial intelligence the problem is to show how the notion of motive differs from the notion of goal, the central notion in the kind of goal-directed reasoning that researchers study in multi-agent systems of deliberation.

In *The Adventure of the Three Students*, exam scripts for a scholarship award had been stolen from the office of a lecturer, and of the students who were to sit the exam, only three had access to the building at that time. From their tutor, Holmes learns that one of them was struggling academically, and also had been known to be a risk taker who was willing to push the boundaries of what is socially acceptable. Both of the others should have been able to do well on their own strength. In a small deviation from the way Doyle tells the story, we can also imagine that the weakest of the students had been overheard saying "If only I could lay my hands on the exam papers, otherwise I'll be finished!" We could say that in this case, all three students had the opportunity to steal the papers, but only one had a sufficiently strong motive to take the risk of detection, shaming and being expelled from university. The other two might have had the same incentive to take the papers as a random person, maybe because of their "market value", selling them to the student who needed them. But over and above the average person, the man who was known to especially desire the papers in advance, having a pronounced emotional attachment in favor of them, had a strong desire to take hold of them. In order to read the questions them, he had to obtain them. And since one way of obtaining the exam papers from the lecturer's office, we draw the conclusion that it is more plausible that this man was guilty of breaking into it, as opposed to the other two. Of course, this is only a small amount of evidence, and indeed, while it is the conclusion drawn initially by the lecturer, is ultimately mistaken.

Examples like this one, where there is a chain of reasoning to the conclusion that this man is the most likely suspect to have taken the papers, are puzzling. His statements provide evidence of a sort, based on a chain of reasoning supporting the conclusion. But what kind of evidence and inferential process are we working with? How can it be identified and defined as a specific kind of evidence? And what is the chain of reasoning it is based on? Does it have a structure? Can it be analyzed as specific type of rational argumentation used to justify a claim?

Maybe the student didn't intend or plan to steal the papers, even though he had a motive to do so, as explained above. He just found himself in the office by chance, but we can easily imagine that once there,

overcome with fear of the exam and desire to gain the scholarship, he succumbed to temptation. He didn't make any long term plan to obtain the papers if he was merely passing by, and saw them on the desk by chance. On this account, he evidently had no long term goal to steal the papers. Still, if he took them, some planning or deliberate action of a sort must have been involved. He had to use some method to get the papers. Perhaps, for example, he had to ascertain that the office was empty, so he could pick them up from the desk. Actions were involved, and he had to do one thing in order to do another. Thus, although it was not his original intention to steal the papers, and it may not have been part of his character to steal anything, we can say that he had a motive to steal them. If he did take them, once he found himself in the office and saw them, presumably his immediate goal was to read them. .

Based on this simple example, we can say that there is a distinction between a motive and intention. We can also say that both motives and intentions are related to planning and to sequences of actions used to carry out a goal. On one reconstruction of what happened in the student's case, his motive, the mainspring of his action, was his desire to pass the test. It was this motive, we hypothesize when reconstructing the actions and goals in the case, that propelled him to take the papers from the desk. On our analysis in the account that follows, a motive will be defined an immediate goal that led (through a sequence of goal-directed reasoning) to an action. We construct the hypothesis that an agent had a motive by reasoning backward from the evidence of the agent's actions and words. The first step in our effort to build a framework to support this definition is to examine the structure of an agent's goal-directed reasoning to an action.

In the story, the lecturer justifies his suspicion by three slightly different points – the academic weakness of the student, his unreliability and risk taking character, and a past incident of dishonesty. For him, they all form part of the everyday notion of “motive” – what “made the student do it”. In a trial setting however, the last two would be problematic and most likely ruled inadmissible. Is it possible to distinguish these different types of reasoning on the basis of a logical analysis alone? Can the distinction be formulated in a way that is clear enough even for a computer to understand it, or is it ultimately reliant on human “intuition”, and with that necessarily problematic, and in every single case contested? In what follows, research and concepts taken from artificial intelligence and agent technology will therefore fulfill a dual task: First, we may be able to profit from the conceptual distinction and analysis that was carried out by this community in their quest to enable software programs (“autonomous agents”) to reason reliably about the plans and motives of other such agents. Second, by looking at examples of concrete implementations of such approaches in computer systems, we hope to partially answer the question to what extent the distinction found in the law are “objective” and repeatable – if a computer can perform the task of distinguishing for example arguments from motives from arguments from character, then the distinctions are definition intersubjective and independent from personal bias or convictions of the lawyers arguing a case.

#### 4. Practical Reasoning

Let's start with the notion of goal-directed agent reasoning. It provides a basis for further investigation of the claim we made in the exam paper example: that motive is different from goal and different from intention and character. Researchers in artificial intelligence, cognitive science and argumentation think of goal-directed reasoning (or practical reasoning) as a kind of reasoning carried out by an agent, of the kind familiar in multi-agent systems. An agent could be a person, but equally a software entity programmed with (or to select) goals, that has the capability of collecting information from its immediate environment, can take or direct actions that will affect that environment and have consequences, and can take incoming information into account to correct its own actions (or simulated actions) to direct them toward the goal.

Agents also have another important characteristic. They can communicate with other agents, and they obtain some of the information they receive and use to direct their actions from other agents. Agents, (often called rational agents) are autonomous, meaning that their actions are not completely determined by the programmer who has designed the agent.

The following scheme represents the simplest and most basic kind of practical inference that is readily familiar to all of us. The first-person pronoun 'I' represents an agent. More correctly, it could be called a rational agent of the kind described by Wooldridge (2000; 2002), an entity that has goals, some (though possibly incomplete) knowledge of its circumstances, and the capability of acting to alter those circumstances and to perceive (some of) the consequences of so acting.



### *Basic Form of Practical Inference*

I have a goal, G.

Carrying out this action A is a means to realize G.

Therefore, I ought (practically speaking) to carry out action A.

This basic form of practical inference is very simple, yet we all recognize its importance as a kind of reasoning we use in daily life.

To show how we have to increase the conceptual vocabulary for our analysis step by step, we introduce yet another of Conan Doyle's stories, *The Adventure of the Noble Bachelor*. Sherlock Holmes's client in this story is the unfortunate Lord St. Simon. On the day of his wedding to a rich American heiress, between the actual wedding ceremony and a reception for the guest the bride suddenly disappears. A suspicious female had been seen talking to the groom, possibly a past love of his but socially inferior and not a candidate for marriage. Has his new wife been kidnapped or murdered? To solve this issue, Lord St. Simon, Watson and Holmes engage in a chain of practical reasoning.

We begin with modifying an example by Atkinson, Bench-Capon and McBurney (2004, p. 88). In this example of a simple and basic kind of practical reasoning, the goal is stated in the first premise as something the agent wants.

I want to be at my wedding reception at 10:30.

If I have a rest until 9.50 and leave the hotel at 10.00, I will be at the reception at 10.30.

So, I shall leave my room at 10.00.

This extremely simple example of practical reasoning shows us why we are dealing with a mystery in the first place. The bride should have had a strong motive to perform a certain action. She did not perform this action, and this needs explaining.

But practical reasoning may involve several complications, even in such a simple case. One is that an agent sometimes has a choice between two or more actions, each of which is a means to the goal. She could have waited longer, but taken a coach to arrive in time. How can I choose the best means? I must compare the alternative means to determine which is best, or would most efficiently realize the goal. Other factors, such as costs and time available may have a bearing, as consequences of the various actions that are being considered. Would it save time if she took the coach? Is one more expensive than the other? A more complex model of practical reasoning called *practical reasoning with several alternative actions* may be useful here. This more complex model of practical reasoning requires specification or development of criteria by which the alternative means can be compared.

Examples such as this one suggest that a simple model of practical reasoning will not suffice—instead a deeper model that takes other factors into account is necessary. Some or all of an agent's multiple goals may have a bearing. Moreover, the model must take into account inconsistencies in an agent's goals. Suppose her goal was to arrive in time. She could have solved the problem by waiting until 10.20, and then running down the street as fast as she could. That action would be too hasty, in light of convenience, her dress, and many other factors. If nearly all of her life was going to be spent in the city, she might buy her own coach and hire a permanent driver, but she also has factors such as style and comfort in mind. Moreover a salesperson might point out that, "This coach uses fewer horses than the other you are considering, is therefore easier to keep in a city, is cheaper to buy, and has style and comfort," suggesting additional goals that could increase the complexity of any model of the decision-making task.

Practical reasoning, in its basic form, is defeasible, meaning that it can be undercut or defeated by asking any one of the following critical questions in a dialogue, or by posing counter-arguments that defeat it. To take all these factors into account, any given instance of practical reasoning needs to be evaluated by the asking of critical questions that pinpoint weak parts of the argument that can be subject to doubt.

### *Critical Questions for Basic Scheme for Practical Reasoning*

(CQ1) What other goals should I consider that might conflict with G?

- (CQ2) Are there alternative actions that would bring about *G* but not *A* that I should consider?  
 (CQ3) Among the action bringing about *A* and these alternative actions, which is arguably the most efficient?  
 (CQ4) What grounds are there for arguing that it is practically possible for me to bring about *A*?  
 (CQ5) What consequences of my bringing about *A* should I also consider?

Modeling such critical questions, and how they interact with the basic scheme for practical reasoning is more complex than the traditional logical method of evaluating an argument as a set of premises and a conclusion. The latter method employs a set of propositions (statements) that are true or false, whereas questions and answers require viewing an argument as a dialogue, a much more complex model. Hence, naturally, a question arises: whether the factors cited in the critical questions could be built in as additional premises of the argumentation scheme. One way to do this is to add several other premises to the basic scheme,

One way to do this is to add several other premises to the basic scheme, taking other relevant factors into account to produce a complex scheme of the following sort (Walton, 1990, p. 112; Girle, Hitchcock, McBurney and Verheij, 2003).

#### *Complex Scheme for Practical Reasoning*

I have a goal, *G*.  
 Carrying out this action *A* is a means to realize *G*.  
 My other goals have been considered.  
 Alternative actions have been considered that are means of achieving *G*.  
*A* is the best means for achieving *G*, of all the actions considered.  
*G* is a goal that is possible to achieve.  
 There are no negative consequences of carrying out *A* that outweigh *G*.  
 Therefore, I ought (practically speaking) to carry out action *A*.

This more complex scheme is based on values. As the fifth premise shows, the agent is picking the best of the available set of actions that seem to be means of achieving the goal. In the seventh premise, the agent is weighing the negative value of some of the consequences against the positive value of the goal. These observations suggest the need for a variant of the scheme for practical reasoning that takes values into account. Consider another inference of the practical reasoning type.

Good manners require that I attend the reception in my honor.  
 If I leave at 10.00 then I will be there in time for the reception  
 So, I shall leave at 10.00.

As Atkinson, Bench-Capon and McBurney (2004, p. 88) pointed out, the action in the conclusion is justified in this inference in relation to the agent's value, showing good manners or observing socially accepted rules. On their account (Atkinson, Bench-Capon and McBurney, 2005), three elements should be considered, the goal the action and the value. They analyze values as social interests that explain why goals are desirable. Values are reasons for holding goals.

This model distinguishes between values and goals. One reason for this approach that in ethical, legal and political deliberations, values can be instrumental in persuading other agents to undertake a course of action. The assumption that all participants share some social values may be important to take into account as a premise. As Atkinson, Bench-Capon and McBurney (2006) have showed, decisions for action can often be based on values that provide reasons to support goals. In such cases, values may be additional relevant factors in practical reasoning that need to be considered.

With all this in place, we can now not only better analyze the *Mystery of the Noble Bachelor*, we are also in a position to shed a clearer light on our initial problem. In the story, Lord St. Simon, the embodiment of conventional morality, assumes not only a shared value system that governs and explains the actions of him, his bride and indeed Holmes and Watson. He assumes that this value system is so strong and non-negotiable that his bride could not possibly have had a motive not to come to her reception. Since due to her values, she cannot have had a motive not to come, some outside force must have prevented her from attending. Both Watson and Inspector Lestrade follow this implicit line of

reasoning, to the point that he latter starts draining a nearby lake in search of her body. Holmes, by contrast, remains unconvinced. By all means a more unconventional person himself, he can envisage situations in which even a bride on her wedding date can have conflicting goals. If the bride furthermore grew up in America – a country known for its unconventional morality, the likelihood that her goals are dictated in part by another system of values raises considerably. On this basis, Holmes quickly solves the mystery: Between the wedding and the reception, something must have happened that gave the bride a new set of goals. Given the strong social values to attend one's own wedding reception, this must have been a considerable event, such as meeting a former lover. This also explains some of the other facts of the case: a stranger caught her bouquet after the wedding, the same stranger was seen giving her a note which later turns out to be a hotel address. Having “abductively” hypothesized the existence of a former lover as the best possible explanation for her change of mind, Holmes uses his resources to quickly locate him in one of the better hotels. It turns out that many years ago, the two not only were lovers, but indeed married. She had believed him dead, only for him to intercept her after the ceremony. Afraid of the scandal, she decided to elope with him without telling anybody of what was, technically, the crime of bigamy.

We note that the story thus contains a small paradox: it is only because *all* the agents in the story - Holmes, Watson, Lestrade, bride and groom - all share a social value system that creates certain goals for them ( not missing, at all costs, one's wedding reception) that we are having a mystery in the first place. Without this *shared* commitment, she would not have minded telling everybody that the wedding was off, and even if she had disappeared without notice, nobody would have been surprised. The stronger the putative motive for attending, the greater the mystery. Holmes however realizes that we sometimes have conflicting goals, and that we sometimes are governed by mutually contradictory value systems. This allows him to explain her absence as voluntary, *even if* she had a strong motive to attend.

Trials, so we argue, always have the structure of the *Noble Bachelor* – and this explains to some extent the paradox that we described in the beginning. A crime has been committed. This means that trivially, someone must have had a motive to take the action that brought it about. This follows through simple application of the reasoning schemes that we have discussed so far, and will find its full justification in the “belief, desire, intention” scheme of human action to be discussed below. This scheme is so deeply rooted in our cognitive approach to make sense of the world that we cannot but draw this conclusion: there is no human action without some motive. Because it is a trivial inference, it does not tell us much. This explains why there is no legal requirement to show motive, the actions speak for themselves, and so to speak carry their motives with them.

At the same time, the very fact that we are in a trial context also means that the action in question not only violated social values, but also that there was a mechanism in place to give strong motivation to adhere to these values –the punishment. The question therefore is not who had a motive to commit this action. Trivially, whoever committed the deed had a motive. Rather, the question is: whose motive was strong enough to cancel out all the other motives he or she had not to commit the crime, from the shared social values to the possible punishment? The more the prosecution asks in punishment, the more it creates an argumentative burden for itself. *If* the consequences for the action should be as severe as the prosecution claims, only a very strong conflicting motive can explain the action. This explains why, despite the lack of legal requirement, prosecutors will always seek to establish a strong motive, and it also explains the failure of the “governing law” models of reasoning with motives mentioned above. It is precisely because we are dealing with deviant behavior that no premise of the form: “Typically, people will do X” can be part of the argument.

To have a model of practical reasoning more adequate to consideration of such value factors, the following scheme (Bench-Capon, 2003) can be adopted.

#### *Scheme for Value-based Practical Reasoning*

I have a goal *G*.

*G* is supported by my set of values, *V*.

Bringing about *A* is necessary (or sufficient) for me to bring about *G*.

Therefore, I should (practically, ought to) bring about *A*.

We can match a comparable set of critical questions to the value-based scheme and use it to evaluate cases of arguments that fit the scheme. In cases where values have no bearing, we need only use the

basic scheme to evaluate practical reasoning. Otherwise, the scheme for value-based practical reasoning may come into play. One can regard the basic scheme as representative of instrumental practical reasoning, while the values-based scheme is more than merely instrumental in nature.

## 5. Hierarchical Goal-Seeking Systems and Plan Recognition

As described above, research in argumentation, cognitive science and artificial intelligence conceives of an agent as an intelligent entity that can do five things. It can collect data (knowledge) about its external circumstances. An agent can bring about changes in these circumstances. It can store the information it has gathered in a knowledge base (or commitment store). The agent can be aware of some of the consequences of its own actions, adding this information to its knowledge base. And it can take actions in light of its observations of these known consequences. These five characteristics are components representing how an agent reasons and acts in a hierarchical goal-seeking system or "HGS" (Walton, 1990, p. 193). In such a system, agents bring about states of the world, have goals, and can realize these goals by virtue of their ability to carry out sequences of actions organized into habitual routines. We can break down a complex action, of kinds we are familiar with, into a set of components for analytical purposes, forming a sequence of actions that can, in turn, be decomposed into goals and actions linked to the goals. If we were to think of the performance of the task at the most mechanical level, output signals drive actuators to carry out specific, directed tasks that the sequence of reasoning determines at higher levels. The hierarchical goal-seeking system would monitor output actuations, and as a result, they become a part of the incoming data stream forming a circular feedback reasoning process. Looking at the system on a higher level, the HGS organizes and exploits a mass of knowledge in order to reach the goals that the input command has set.

Each step in a hierarchical sequence of reasoning in the system fits the fundamental form of practical reasoning (FFPR) for an agent in an HGS.

### *FFPR for an Agent in an HGS*

Goal Premise: Agent  $\alpha$  has goal G.

Means Premise: Agent  $\alpha$  knows that that action  $\mu$  is a means for bringing about goal G.

Conclusion: Agent  $\alpha$  should carry out action  $\mu$ .

Some explanations of the expressions used here are necessary. A goal is an internal state of an agent associated with a proposition that states the goal. Goals can be at varying levels of abstraction. For example my particular goal at this moment may be to pull a trigger. At a higher level of abstraction my goal might be to fire a shot at John. At an even higher level my goal would be to inherit his money, or to prevent a danger to national security. Generally speaking, an agent's goal can be any proposition that the agent reasonably thinks it can bring about by taking actions 2

Each step in the sequence of reasoning is linked to its neighboring step by a binary relation, which action theory calls a by-relation (Goldman, 1970, chapter 2). The by-relation links different individual act-descriptions into a chain of complex act-sequences. Scholarship in the theory of action regards four kinds of by-relations as centrally important (Walton, 1990, pp. 200-201).

(1) *Causal Generation*. If an agent brings about one state of affairs, and this action causes another state of affairs to occur, we say that the action of carrying out the first state causally generates the agent's action of bringing about the second state. For example, in *the Noble Bachelor* a bystander reaches out to catch the bride's bouquet. Then we may conclude in some situations that the agent carried out the act of catching the bouquet.

(2) *Conventional Generation*. According to Goldman (1970, p. 25), "[c]onventional generation is characterized by the existence of rules, conventions, or social practices in virtue of which an act . . . can be ascribed to an agent . . . , given his performance of another act . . .". For example, when the priest said: "I hereby declare you man and wife" during the wedding we may conclude that the agent performed a marriage.

(3) *Simple Generation* (Circumstantial Generation). If an agent carries out one action, from what we know of its circumstances, we may infer that the agent has also carried out another action. For example, when the bride excused herself pretending to have a headache, we can infer that she was also planning her escape.

(4) *Augmentation Generation*. If an agent carries out one action, then its so doing can, in some cases, be augmented by further information. For example, when the unknown stranger gave the bride a piece of paper, we may infer from what we know of the circumstances that he gave her a piece of paper with his address on it.

The generation of actions at each level supports a sequence of reasoning that can be used to justify or explain the drawing of an inference between a pair of act-descriptions. This means that gaps in the chain of reasoning are filled in by applying the principles stated in the four types of generation specified above. However, the reliability of each inference is relative to what is known about the particular circumstances of the action, and so the sequence of reasoning is, in typical cases, not deductively valid. Instead, the inference is a defeasible kind of reasoning based on common knowledge about normal and expected ways of doing things with which a group of agents is familiar.

Using the by-relation, analysts can connect actions to other actions of the same agent in a sequence, forming chains of reasoning in the process. One kind of sequence is a “forward chaining” kind of practical reasoning from an agent’s goals and knowledge of its circumstances to a conclusion about an action. In addition, however, we can use such sequences to infer the agent’s goal, its reason for the action, from the agent’s action – the “backward chaining” aspect of the inference. One can model this kind of reasoning with syntactical strings that represent the by-relation, where  $\alpha$  is an agent and  $\mu$  represents a state of affairs (proposition) brought about by an agent: “bring about  $\mu \rightarrow$  bring about G.” Such a syntactical string has the semantic meaning that if the agent brings about  $\mu$ , it can be inferred as a hypothesis that the agent had goal G as part of its reason for bringing about  $\mu$ . Using this syntax, we can represent the structure of forward practical reasoning from goal to action as follows.

*Forward Sequence of Practical Reasoning to an Action*

$Com_{\alpha} ((GOAL\ G) \ \& \ Know_{\alpha} (\text{bring about } \mu \rightarrow \text{bring about } G)) \rightarrow Com_{\alpha} (\text{bring about } \mu)$

The  $\rightarrow$  represents a defeasible inference from a set of premises to a conclusion about an action. The *Com* operator represents the agent’s commitments at any local point in a dialogue. If the agent is committed to a goal, as indicated by the evidence of its actions and utterances in a dialogue, and also committed to some means it knows will bring about the goal, then as a rational agent it must also be committed to the act of carrying out these means. (The ‘must’ does not refer to logical necessity, but only to a kind of defeasible practical reasoning, meaning that if the agent is committed to the premises, it should--practically speaking-- carry out the action expressed in the conclusion, according to the constraints of the plan it has adopted.) There is also a backward sequence of practical reasoning that corresponds to each instance of the forward sequence.

*Backward Sequence of Practical Reasoning from an Action*

$(\alpha (\text{bring about } \mu) \ \& \ Know_{\alpha} (\text{bring about } \mu \rightarrow \text{bring about } G)) \rightarrow Com_{\alpha} (GOAL\ G)$

Suppose that we have evidence that an agent has brought about a particular action, and evidence that it knows, based on its plan library<sup>3</sup> and its knowledge of its circumstances, that bringing about this action is the means to bringing about a certain goal. Based on such evidence, we can infer by defeasible reasoning (i.e., subject to revision as new evidence comes in) that the agent is committed to this goal.

Researchers in artificial intelligence have done considerable work on what is called *plan recognition*. Some systems use top-down plan recognition by drawing inferences from an agent’s expected goals to its actions and beliefs. Others use bottom-up inferencing from the agent’s observed actions to construct a plausible plan that can be attributed to the agent and account for its expected goals and beliefs. Carberry (1990 chapter 2) presented a survey of several such systems developed up to 1990. Her outline offers a useful introduction to how the technology of plan recognition works. She noted (p. 74) that experimental evidence of natural language conversational interactions suggests that people infer goals of other agents from what they say and do, and expect other agents to infer their own goals from the same kind of evidence, and use them in addressing their needs. Most relevant for our purposes, she argued that plan recognition must play an important role not only in designing such mechanisms but also in understanding natural language dialogue.

In our analysis of motive evidence, and following Carberry, motives are immediate goals to which an agent is strongly committed and has adopted as a mainspring of an action it has taken. One agent can infer that another agent, *B*, with whom it is engaged in a dialogue has a particular motive by using the evidence that the dialogue has produced about *B*'s statements and actions. To this point, we have analyzed the reasoning through which actors can infer motives from evidence on the plan recognition model. Development of a systematic structure for evaluating such evidence is a greater challenge.

There are two philosophical approaches to the problem of evaluating the strength of practical reasoning. According to the commitment-based theory (Walton, 1989), two agents (in the simplest case) interact with each other in a dialogue in which each contributes speech acts.<sup>2</sup> On the commitment-based approach, one can conceptualize practical reasoning in a dialogue format using one of the argumentation schemes above along with the set of critical questions matching the scheme. The initial argument is strengthened if critical questions are answered, or is undercut or defeated if they are not.

The alternative cognitive model, which rests on the notion of an agent updating its set of beliefs based on incoming information, is the BDI (belief-desire-intention) theory. According to the BDI theory (Bratman, 1987; Bratman, Israel and Pollack, 1988; Wooldridge, 2002; Paglieri and Castelfranchi, 2005), an agent has a set of beliefs that are constantly being updated by sensory input from its environment. The agent builds up desires from these beliefs, and wants that it then evaluates (by desirability and achievability) to form intentions. An intention is persistent goal that is not easily given up. In Bratman's (1987) version of the BDI model, forming an intention is comparable to adopting a plan that includes desires (wants) and beliefs. Pollock (1995) but added what he called "likings", as well as desires, that need to work in combinations with beliefs and intentions.

The two models are different insofar as goals and commitments differ from desires and beliefs. The latter are deeply psychological notions whereas the former are procedural notions based on rules governing rational argumentation between two parties. Nevertheless, many influential writers on practical reasoning tend to blend the two models together. Searle (2001) purports to offer a BDI model of practical reasoning, but, like Bratman, often shifts to the language of commitment. This can be confusing, because acceptance, a term that can be taken as equivalent to commitment, does not imply belief. Bratman, Israel and Pollack (1988, p. 347) wrote, "The fundamental observation of our approach is that a rational agent is committed to doing what she plans". This wording uses the language of commitment, although Bratman, Israel and Pollack officially see themselves as promoting the BDI model. Despite much blurring of the borderlines in the literature, the two approaches are, in principle, distinct. The commitment model uses argumentation schemes, expressed in a dialogue (question-answer) format in which two agents reason together. The BDI model uses intentions, beliefs, likings and desires that represent psychological states of an individual decision-maker that are modified as the agent acquires new data from its sensors, or other sources of information.

## 6. Motive as Emotion Inferred from the Evidence of Fact

Wigmore (1931, p. 146) defined motive as a specific emotion or passion that is likely to lead to the doing of a particular act, for example desire for money to the active fact of robbery, or angry hostility to an act of violence. While not necessarily a definition Doyle would have subscribed to, it gives us a good starting point to link legal analysis and the analysis of our detective. In law, the fact-finder may not attribute a motive to one party or the other without proof.

According to Wigmore (1931, 146), evidence of motive is usually circumstantial, consisting of the conduct of the person and the events in the particular situation tending to excite the emotion. He added (p. 147) that there is an unfortunate ambiguity in the word 'motive'. To illustrate this ambiguity, he cited an example in which a defendant is accused of burning down the plaintiff's house. Plaintiff contends that the defendant's motive for burning down the plaintiff's house is the plaintiff's prior prosecution of a lawsuit against the defendant. The ambiguity in "motive" is that writers might sometimes use the external fact of the plaintiff's lawsuit as the motive. But they might also, more properly, refer to the defendant's hostile and vindictive emotion about the lawsuit as the motive. On Wigmore's analysis, the term motive may have

<sup>2</sup> Each has a commitment set, and as the one asks questions that the other answers, commitments are inserted into or retracted from each set, depending on the type of move (speech act) each speaker makes. A commitment is a proposition that an agent has gone on record as accepting (Hamblin, 1970; 1971). One type of speech act is the putting forward of an argument. When the one agent puts forward an argument, the other can reply, either by asking critical questions or by putting forward a counter-argument.

either meaning. It can be a particular event or fact known to an agent, but it can also be the lively emotion of that agent with respect to event.

According to Wigmore (p. 147), there is always a double evidential step involved in proof of a motive, used as evidence of another fact. When it is said, in the example above, that the plaintiff's prior lawsuit is evidence that the defendant had a motive for burning down the plaintiff's house, we might think that we are making an inference directly from the prior lawsuit as evidentiary fact to the burning as the proposition proved. But, according to Wigmore, there are really two steps involved, one inference from the lawsuit to an emotion and a second from the emotion to the act. What Wigmore describes is a complex process of reasoning, from facts in evidence to the hypothesis of a motive, then from that motive, along with circumstantial evidence again, to the inference that a particular agent committed an action. What model of reasoning could be used to exhibit this complex process of reasoning?

The entry point into the study of motive is the contrast between the model of the rational person as agent and the model of the psychological person, driven by desires and emotions. The model of the rational person is the commitment-based inference model that is characteristic of research on practical reasoning. On the forward chaining variant of this model, an agent has a goal, sees a means of carrying out this goal, and then draws a conclusion that it should carry out the action corresponding to the means. The backward chaining variant is a practical inference in the opposite direction. A second agent, a critic or evaluator of the action, sees the first agent perform the action, and then judging from what it knows about the first agent and its circumstances, draws a conclusion about the presumed goal that the first agent had in mind when it performed the action. The model of the psychological person parallels the commitment model of the rational person, except that desires or wants, rather than goals, drive actions in the psychological model. A desire or want is like a goal, except that it is not simply a proposition, but an attitude about something representing an excited emotion for or against something. On the forward chaining variant of the psychological person model the excitement of a strong desire drives an agent towards carrying out an action. On the backward chaining variant, the second agent, a critical or evaluator of the action sees the first agent perform the action, and then judging from what it knows about the first agent and its circumstances, draws a conclusion about the first agent's motive for performing the action.

The concept of motive in criminal law involves an even more sophisticated form of reasoning, a variant on the backward chaining variant of the psychological person model. In a typical example, there are a number of suspects who might have committed a crime. They might all have had an opportunity to commit the crime. But one of them could have had a motive to commit the crime that the others lacked. For example, in a case that Leonard (2001, p. 442), considers, a nurse was charged of theft of the controlled substance Demerol, a painkiller, from a locked hospital cabinet. Four other nurses also had access to the cabinet, but all denied taking the painkiller. The defendant admitted, however, that she had once been addicted to Demerol. A urine test result showed that she had used Demerol recently. There was evidence as well that four years earlier, she had stolen Demerol from a hospital where she worked, and that her nurse's license had been suspended for that reason. The court initially excluded this evidence of earlier conviction, reasoning that a person's prior conduct would be inadmissible for the purpose of showing a propensity to act in accord with that conduct. An appeal court overturned this earlier finding, ruling that the basis of the conviction was not propensity or character evidence, but that being a Demerol addict gave this person a motive to tamper with Demerol-filled syringes. The appeal court argued that nobody was suggesting that any of the five nurses might have wanted to steal Demerol in order to resell it, but rather that this particular nurse's addiction to Demerol gave her a motive for consuming it. Thus the court essentially argued that she stole it in order to fulfill her desire to consume it, and that this evidence of the motive to consume it was admissible to show that she stole the Demerol.

The court employed a kind of backward chaining of reasoning using the model of the psychological person. Any of the nurses could have stolen the Demerol. What distinguished the defendant nurse from the other four was the evidence that she had a motive that the others lacked: her addiction to Demerol. Indeed, she had even been previously convicted for having stolen Demerol from a hospital in the past. The prosecution used this evidence to show that defendant had a desire to consume Demerol. The existence of that desire was the motive that the prosecution attributed to this agent, the nurse. The argument was quite convincing. It would be inappropriate to say that her goal was to consume Demerol. While that description could be roughly accurate, it would attribute too much rationality to her action, that is, it would entirely account for her action with the rational person model. The psychological model would better explain her action by attributing to her a motive to consume Demerol. That motive would not be a goal, strictly speaking, but a strong emotion or drive, linked to the action of stealing Demerol from the

hospital. When the court looked at all the facts of the case (essentially using the backward chaining variant of the psychological model) it not only concluded that that the evidence would be admissible to show such a motive, but then sanctioned the use of forward chaining to infer the action from this motive, given the opportunity to steal the Demerol from the hospital. Finally, the court used another kind of reasoning to look at the different hypotheses on the question of who, among the five suspects being considered, was the most likely person to have carried out the theft. Since this particular nurse alone had this motive, the court concluded that the prosecution had offered sufficient evidence that she was the one had stolen the Demerol from the hospital.

## 7. Motive as Evidence

According to Leonard (2001, p. 447), citing Wigmore's discussion of motives (see above), proof of motive is typically indirect, from circumstantial evidence.<sup>3</sup> In other words, triers of fact infer motives from factual data, such as the actions of agents, that witnesses can observe. In turn, on Leonard's analysis, fact finders would characteristically rely on motive as part of an inferential sequence of reasoning that has two steps. The first step in that sequence is the inference from the existence of the motive to the observed actions or behavior of a person on a particular occasion. The second step can lead to one of three facts (Leonard, 2001, p. 447):

- (1) that the person is the one who committed the act in question,
- (2) that the act in question occurred,
- (3) and that the actor behaved with the required state of mind (in criminal cases).

In an example from Leonard (2001, pp. 447-448), the defendant was charged with the murder of the victim, but claims not to have been involved. The prosecution has evidence that the defendant had been involved in a car theft prior to the killing and that the victim had learned about the theft. Also, the victim had threatened to reveal the theft to the police. In this case, evidence of the theft could be admissible to prove the defendant's motive, and it could provide a basis for drawing an inference about the defendant's possible behavior. The way Leonard structures the inference in this kind of case (p. 448) is a paradigm example of how inferences based on motives can be part of a chain of reasoning used to resolve questions of fact.

**EVIDENCE:** Defendant stole a car, victim was aware of the fact, and victim threatened to inform the police.

**INFERENCE:** Defendant had a motive to prevent victim from revealing the theft to the police.

**CONCLUSION:** Defendant murdered victim to prevent victim from revealing the theft to the police.

We see the reasoning in this kind of case as typical of abductive reasoning, or inference to the best explanation—a kind of reasoning that characteristically infers a conclusion from a set of observed or given facts or data. Typically, in such a case, the two sides have presented two opposed accounts, or stories, at trial. Leonard's reconstruction, above, nicely models the structure of the argumentation that the prosecution employed in the argument. The prosecution's explanation, as shown in Leonard's analysis of its inferential structure above, takes the form of an argument from motive. In response, the defense would usually offer its own account. For example, the defense might argue that the victim died at the hands of another person, or died of natural causes. Each such account would offer an explanation of the facts.

<sup>3</sup> This assertion is overstated because we can imagine a situation in which the defendant made statements before the crime, or at the time of the crime, indicating a motive. Hence we used the qualifier 'typically'.



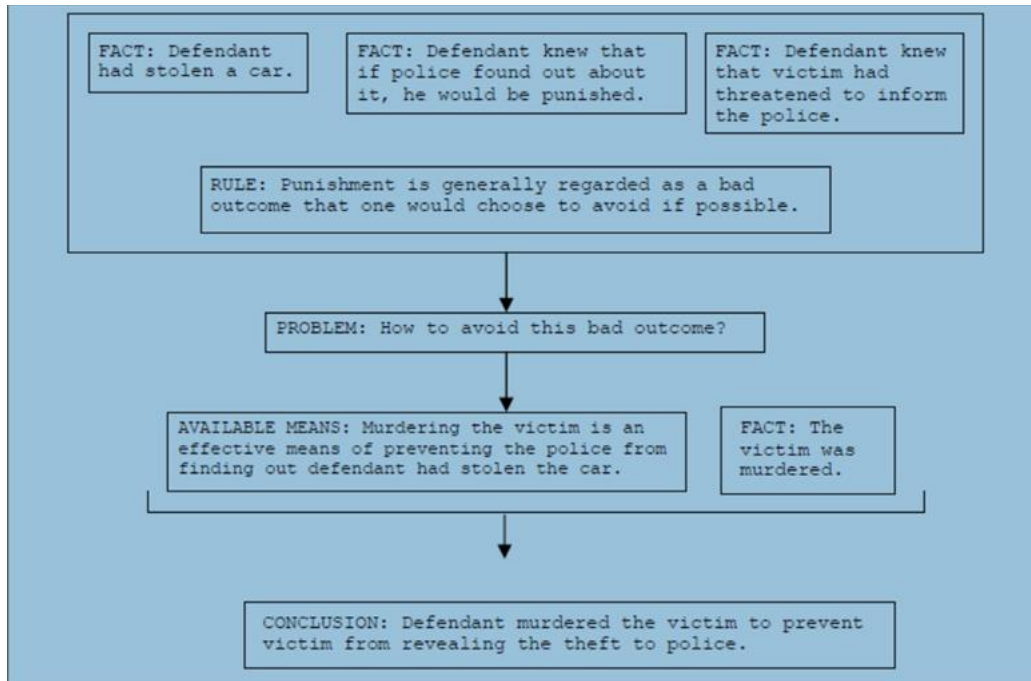


Figure 1: Sequence of Practical Reasoning in Car Theft Case

First of all, there are three items of evidence cited as facts at the top of figure 1. We add to this group of propositions an implicit premise in the form of a common knowledge generalization.<sup>4</sup> This proposition is stated as a rule to the effect that punishment is generally regarded as a bad outcome that one would choose to avoid if possible. Once we put all these four propositions together as a database they form what we would recognize as a problem for the defendant: if possible he needs to avoid this bad outcome of being punished.

We must draw on resources of common knowledge again to fill in another implicit premise in the sequence of reasoning. We all know, and the defendant would know, that murdering the victim would be an effective means of preventing him from informing the police that the defendant stole the car. So at this point in the diagram, we draw an arrow from the problem box to this available means premise box. We, as evidence evaluators, are reconstructing the sequence of practical reasoning in the case based on the facts of the case and on what we know as common knowledge about the normal and expected ways of carrying out actions. On this basis, we can now reason backwards to construct a hypothesis about what the defendant did. Of course, this is merely a hypothetical reconstruction, representing an account or explanation that the prosecution might use to explain how defendant could be guilty. It is no more than a plausible explanation of what the defendant might have done, or plausibly did, given the facts of the case as we know them.

At this point in our model of reasoning we enter a new premise, the fact that the victim was murdered. From this fact, taken together with what we know about the prior sequence of reasoning offering an account of the defendant's practical reasoning, a trier whose reasoning followed the diagram would draw the conclusion that the defendant murdered the victim to prevent the victim from revealing the theft to police. This proposition is represented as the ultimate conclusion pictured at the bottom of figure 1. Hence in this case, the reconstruction of what happened, based on the facts, exhibits a forward and backward sequence. At the top of the diagram the reasoning goes forward, representing what is presumably the solution the defendant arrived at to solve his problem. But at that point, based on the evidence that the available means, and thereby reaching the conclusion in the form of a hypothesis or explanation of the previous sequence of steps reconstructed. The hypothesis is the proposition that the defendant murdered the victim to prevent the victim from revealing the theft to the police.

<sup>4</sup> Defined more fully below.

## 8. Abductive Reasoning

The notion of abductive reasoning was a latecomer in the history of logic, but it has become a useful tool in artificial intelligence to model defeasible reasoning (Josephson and Josephson, 1994). The abductive model also applies very well to trace reasoning in legal fact-finding. For example, if a footprint matching the shoe of the suspect is found at the crime scene, evidence of the trace would be admissible to show guilt. Using abductive reasoning, we can infer an explanation for the shoeprint that the forensic investigators found at the scene, and the discovery that defendant owned a matching shoe: that the suspect, wearing the shoe, left the imprint at the scene when he was there. Abductive reasoning is different from deductive and inductive reasoning, because it is based on explanations of given appearances, and on conclusions drawn only tentatively. The reasoner may abandon those conclusions if unable to answer critical questions appropriately.

It is highly controversial whether there are only three kinds of logical reasoning, deductive, inductive and abductive. The originator of the phrase 'abductive reasoning' was C. S. Peirce, who advocated this tripartite classification of reasoning. Some defining features of each type of reasoning can be illustrated by using Peirce's original example (described by Preyer and Mans, 1999, p. 12).

### *The Canonical Example of Abductive Reasoning*

**Deductive Reasoning:** Suppose a bag contains only red marbles, and you take one out. You may infer by deductive reasoning that the marble is red.

**Inductive Reasoning:** Suppose you do not know the color of the marbles in the bag, and you take one out and it is red. You may infer by inductive reasoning that all the marbles in the bag are red.<sup>5</sup>

**Abductive Reasoning:** Suppose you find a red marble in the vicinity of a bag of red marbles. You may infer by abductive reasoning that the marble is from the bag.

Many would argue that abductive reasoning is a species of inductive reasoning<sup>6</sup>, but Peirce (1992, p. 142) wrote, "There is no probability about it. It is a mere suggestion which we tentatively adopt." Peirce also used the terms 'hypothesis' and 'best explanation' in describing abductive reasoning, as shown below, suggesting that abductive reasoning, on his view was quite different from inductive reasoning.

Abductive reasoning is a kind of guessing by a process of forming a plausible hypothesis that explains a given set of facts or data. As Preyer and Mans (1999, p. 12) point out, in this case the hypothesis, 'The marble is from the bag' could "serve as part of the explanation for the fact that a red marble lies on the floor". This account gives a clue about the nature of abductive reasoning, as being a distinctive kind of reasoning in itself, different from deductive and inductive reasoning. Consider the example, and how the conclusion is derived from the given data. I see the red marble on the floor. I see that it is near the bag. I know that the bag contains red marbles. I then construct the hypothesis, or guess, that the red marble on the floor came from the bag. How? Well, the red marble didn't just appear on the floor. It came from somewhere. There is no other obvious source, let's say. Although there is no hard evidence it came from the bag, that hypothesis appears to be the only plausible explanation that offers itself. No other hypotheses are more plausible. The explanation concerns the source of the marble. It could have gotten where it is by coming out of the bag, and somehow (we do not know how) arriving at its present location on the floor. What is significant in the given case is not only the known facts, but also the boundaries of what is known. There is the bare room, the bag of red marbles, and the single red marble on the floor near the bag. No other relevant facts of the case are known. From this set of data, one explanation of the given location of the marble stands out.

<sup>5</sup> The Preyer-Mans example is a little misleading in that, according to standard statistical methodology, drawing an inductive inference from one finding would be too small a sample. To make the example more realistic, consider a finding of several red marbles with no other colors of marbles found in the sample.

<sup>6</sup> As Josephson and Josephson (1994, p. 270) pointed out, the importance of going beyond probabilities for analyzing abductive reasoning has only been recognized by a minority of researchers in artificial intelligence.

As a species of inference to the best explanation, abductive inference can be defined as having three stages. First, it begins from a set of premises that report observed findings or facts – the known evidence in a given case. Second, it searches among various explanations that can be given for these facts. Third, it selects out the so-called “best” explanation, and draws a conclusion that this explanation is acceptable as a hypothesis. The sequence of reasoning in the red marble case could be represented schematically as follows.

Positive Data: the red marble is on the floor, near the bag of red marbles.

Hypothesis: the red marble came from the bag.

Negative Data: No other relevant facts suggest any other plausible hypothesis that would explain where the red marble came from.

Conclusion: The hypothesis that the red marble came from the bag is the best guess.

The best guess is just an assumption, or defeasible conclusion. It could be overturned by new information that suggests otherwise. But given what is known and what is not known about the facts of the case, that hypothesis is the best guess, or the most plausible one. There are lots of other possible explanations. Somebody could have put the marble there to make it appear that it came from the bag, for example. But in the absence of any relevant known facts of this sort, the hypothesis that the marble came from the bag is the only explanation that is given any plausibility by the actual facts of the case. Abductive inference is defeasible, meaning that the conclusion is only a hypothesis that is subject to retraction if further investigation of the facts in the case shows that another of the alternative explanations is better.

## 9. Reconstruction of the Evidence in the Car Theft Case

Reconstruction of reasoning about evidence in the car theft case involves both practical reasoning and abduction.. The diagram in figure 1 already illustrates, in a rough way, how we can combine forward practical reasoning with a backward use of practical reasoning of an abductive kind—we typically use both to arrive at a conclusion about motive. Practical reasoning enables us to reason forward to identify a likely problem for the defendant, representing a hypothesis about how his reasoning about actions might have gone. Then we reasoned backward by abduction to a hypothesis about how he might have solved this problem by taking a certain action. That abductive reasoning relied on the initial forward use of practical reasoning because it utilized as a premise the available means in a situation, and combined this proposition with the fact that the victim was killed. Taken together, those two propositions functioned as premises in an abductive inference that led to the hypothesis that the defendant murdered the victim to prevent the victim from revealing the theft to police.

First, let us exhibit the structure of the forward sequence of reasoning, using an argument diagram comparable to the diagram in figure 1, except that in this new system, the conclusion appears at the top of the diagram.<sup>7</sup> In figure 2, the conclusion shown in the darkened box with dotted outlines at the top is an implicit argument where an argumentation scheme has been applied to premises and a conclusion. The scheme for practical reasoning (of the basic type) and the scheme value-based practical reasoning are shown to apply to each of these arguments.

<sup>7</sup> The diagramming tool we will use is Araucaria (Reed and Rowe, 2005). The user loads a text document into Araucaria where it appears in the left window, highlights each premise or conclusion and then moves them to a right window. Once all the propositions in the argument have been represented, the user then draws an arrow from each premise, or set of premises, to its conclusion. The result is an argument diagram, representing an analysis of the argument. Argumentation schemes, including ones for practical reasoning, can be represented in an *Araucaria* diagram by selecting the scheme from a set. Implicit premises or conclusions can also be inserted.

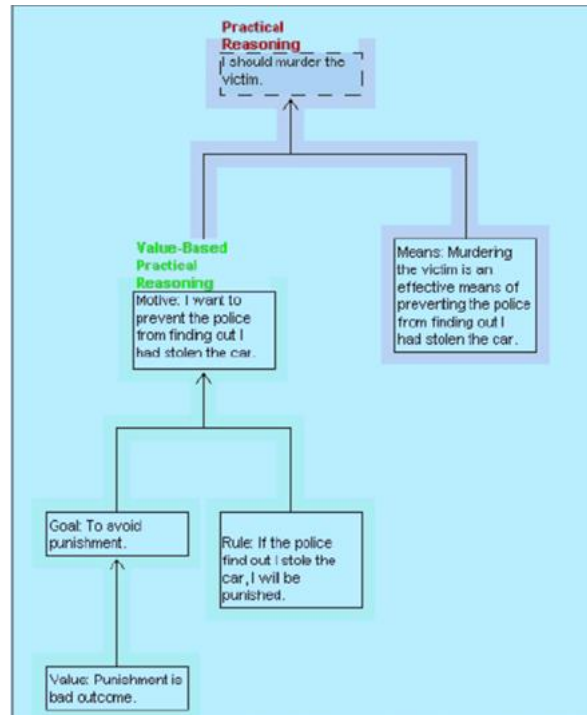


Figure 2: Forward Practical Reasoning in the Car Theft Case

The analysis pictured in figure 2 exhibits the sequence of practical reasoning attributed to the defendant when we reconstruct what we presumed to be his thinking, revealing his motive. Practical reasoning connects the motive and the means to the conclusion that our analysis leads us to believe that the defendant reached. Evidence tending to establish the motive is displayed on the left side of the diagram at the bottom. Here we assume that the defendant's goal would have been to avoid punishment, and combine that with a common sense generalization that if the police were informed about the theft of the car, the defendant would be punished. This part of the argument is an instance of value-based practical reasoning, on the assumption that the defendant would classify punishment as a bad outcome. Figure 2 then displays our analysis of defendant's motive and how it could have linked up with his actions, given that we can fill in steps in the reasoning based on common knowledge assumptions.

The next step is to show how the backward argumentation scheme for abductive inference can combine with practical reasoning. Figure 3 is very helpful.

Figure 3 relies on backward abductive inference of a kind that chooses among several accounts of the data that the parties have submitted in the case, to reconstruct the sequence of reasoning that follows the process depicted in Figure 2. The car theft case is a typical one, in which there could be several suspects with an account or story suggesting each suspect's supposed connection to the crime. Each of those accounts might be equally plausible as an explanation of the event that occurred, namely the killing. One of the accounts, might, however, stand out as a more plausible explanation than the others, in light of one particular suspect's motive for committing the crime. The defendant in our illustration knew that the victim witnessed a car theft that defendant committed. When we employ practical reasoning to reconstruct the whole story, it offers a plausible explanation of why the defendant might have murdered the victim. Figure 3 shows that the argumentation in this case is an instance of backward abductive inference, because the reconstruction of the motive was part of an account that offered a best explanation of the data.

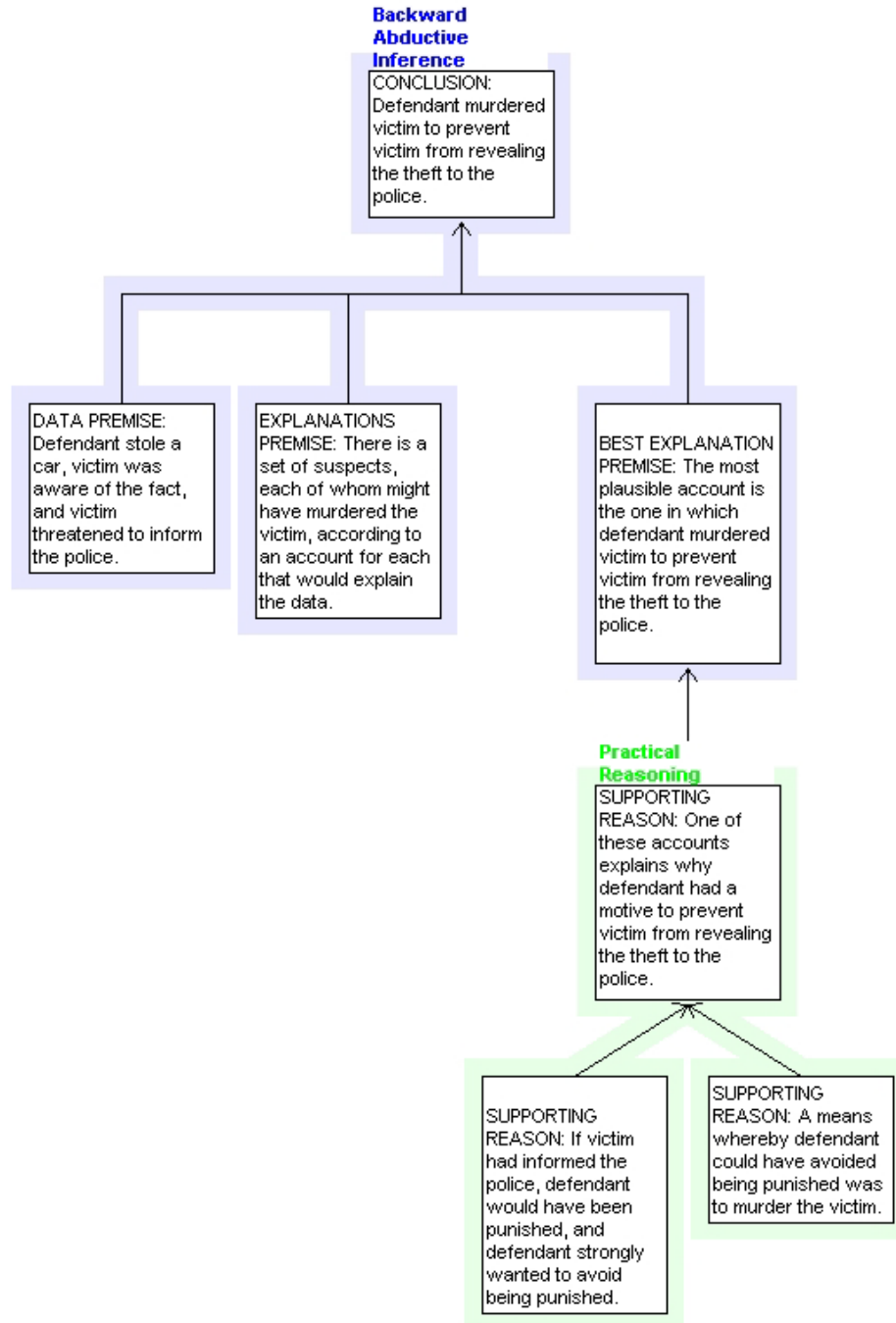


Figure 3: Use of Backward Abductive Inference in the Car Theft Case

We can go even a step further in the analysis of the car theft case as an example of motive evidence, through reconstructing the example as an instance of the forward argumentation scheme for abductive inference. Such a scheme has two components. One is the set of facts in the case. The other is a set of argument diagrams, each one representing a set of premises and plausible conditionals (generalizations) that function as missing parts of enthymemes. In the car theft case, the implicit premises missing parts of enthymemes are assumptions based on common knowledge. Here are the two sets of propositions in the car theft case.

Fact: Defendant had stolen a car.  
 Fact: Defendant knew that if police found out about his having stolen the car, he would be punished.  
 Fact: Defendant knew that victim had threatened to inform the police.  
 From these facts, a motive can be drawn as a conclusion.  
 Motive: Defendant wanted to prevent the police from finding out he had stolen the car.

Figure 4 illustrates the process of drawing a conclusion about the motive. Basically a decision maker forms a set of hypotheses from the facts in the case by drawing on a set of competing accounts, each of which explain the set of facts. The illustration embeds each account in an argument diagram that supports the hypothesis associated with the account. One can evaluate the plausibility of each hypothesis by comparatively evaluating the evidence that each argument diagram sets out. Figure 4 displays a simplified model of this sequence of reasoning, in which two hypotheses are shown. The one on the left includes the argument that the suspect has a motive, while the one on the right is comparable, except that it lacks evidence of a motive.

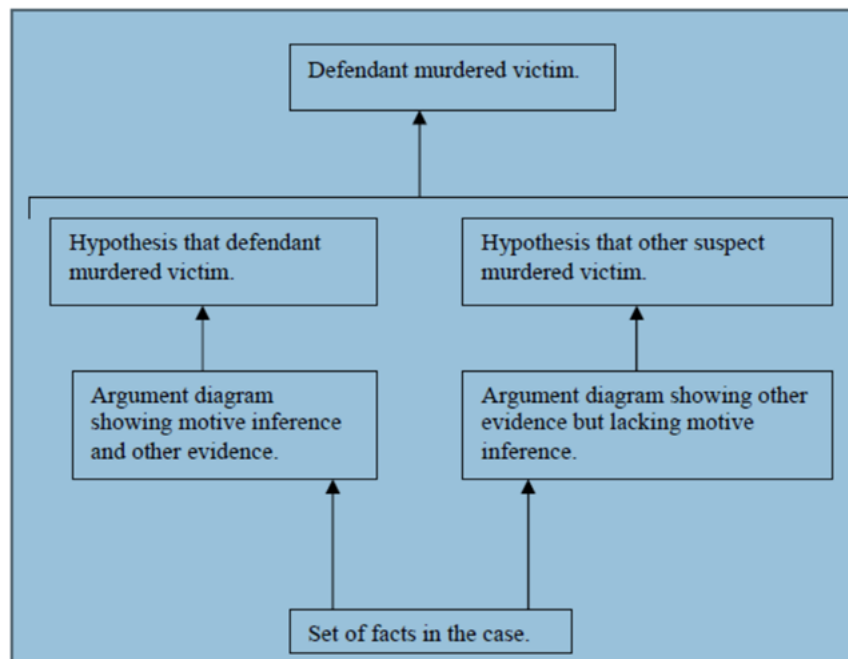


Figure 4: Use of Forward Abductive Inference in the Car Theft Case

The argumentation scheme that applies to the sequence of reasoning in figure 4 is the scheme for forward abductive inference. The argument diagram to which the box on the left would display, if space permitted, would be the argument diagram for motive evidence from figure 2. That part of the diagram would show the forward practical reasoning in the car theft case, which would be part of the account making up the overall diagram. The part of the diagram that involves figure 2 would reflect a theory about the motive, and how it is connected by practical reasoning with the facts of the case.

## 10. Conclusions

The challenge of evaluating evidence for a detective or the participants in a criminal trial (whether in fact or fiction) is not limited to appraisal of trace evidence, wounds and other physical evidence. In addition, one must reach conclusions about how rational human agents, in the intentional pursuit of goals, might affect the world around them in such ways. As a simple example, we might explain the presence of ash as the result of the actions of a human smoker. Philosophers of action have shown that there is more

to an action than a mere bodily movement. Action has a mental component, which, as section 1 points out, is the most important part of the evidential picture for the detective. A fuller explanation of ash evidence might, for example, involve the development of an account in which the perpetrator was waiting for his victim and smoked a cigarette to kill time and fight his nervousness. Julian Barnes' *Arthur & George* involved an even more complex case, involving an inference from the established evidence of the letters and the mutilated pony to a hypothesis about George's motive. On our analysis, we can use a backward sequence of practical reasoning from the facts of the case to formulate a hypothesis about possible motives. A model of that process as a chaining of logical inferences would require insertion of evidence of the agents' actions and statements as data into a plan library, on which we could rely to reconstruct a sequence of practical reasoning. We can then, through an inference to the best explanation, employ the chain of reasoning to support a hypothesis about the motive that likely led to the the particular action at issue.

The problem has been that philosophers and lawyers have lacked analytical tools that they could use to model this mental component, and thereby to reveal the logical structure of motive evidence, in contrast with other kinds of evidence important in legal fact-finding. Recent developments in AI and argumentation, as discussed above, have provided new tools for overcoming this deficiency. Foremost among these has been plan recognition technology, which provides a framework for hypothesizing a rational agent's possible plans from information about that agent's verbal interactions with another. That technology offers a good foundational basis for solving hard problems in evidence law, such as the perennial one of trying to draw clear and precise distinctions between motive and character on the one hand, and motive and intention on the other.

In this paper, having adopted the technology of plan recognition systems such as TRACK, we have followed a model of planning in which an agent has goals, as well as a set of beliefs or commitments, and on the basis of its goals and beliefs, carries out actions. In our theory, the agent uses practical reasoning to make inferences in order to reason from goals and commitments to a decision to carry out an action. In line with that commitment model, we analogized making inferences about motives from the evidence at a trial to the mental processes in which two agents would engage when involved in a dialogue in which each made utterances, in the form of speech acts, like the speech act of putting forward an argument. As such a dialogue proceeds, the agents make commitments to various propositions, according to rules governing the speech acts, and how each speech act, when performed in a dialogue, incurs commitments for the agent who made it. Many scholars of argumentation theory (Gordon and Karcapilidis, 1997; Reed and Norman, 2003; Verheij, 2003; Prakken, 2001; Bench-Capon, 2003a).have already written about formal models of such dialogues, and their import for the analysis of legal argumentation

The main point of our paper is that agent's deliberations with each other using practical reasoning, offer good examples of a process in which one agent tries to determine what the likely goals of the other agent are. To outline that process, we adopted a bottom-up model of plan recognition using backwards practical reasoning. Essentially, the one agent analyzes the statements and actions of another agent, along with known facts of the case, and by backwards abductive reasoning constructs a hypothesis that offers a best explanation of the set of facts. By this means, the one agent reconstructs what it takes to be the motive of the other agent, through drawing an inference from facts and commitments at a given point in a dialogue or investigation to a conclusion that describes the presumed motive of the other agent.

We think that this theory can throw light on problems in evidence law by helping us to better understand the structure of the process of logical reasoning used to draw an inference from a given set of facts to a conclusion about motive. We have used the term 'goal' in a logical sense, making it a very neutral term. A goal can be simply any proposition that is part of a plan linking it to an agent's commitments and actions by the argumentation scheme for practical reasoning. What then is the difference between a goal and intention? Also, what is the difference between an intention and a motive? In our theory, an intention is a goal that fits into the structure of the BDI model, while we do not conceive of a goal, in psychological terms of wants or desires. The difference pertains to evaluation of practical reasoning. It depends on how you want to evaluate practical reasoning in a given case. You can evaluate it on a commitment model or on a BDI model. Our theory of motive evidence can be used in either model. We also think that this theory can be helpful for artificial intelligence by enabling it to develop more sophisticated technology to be used as legal reasoning assistants for various purposes (Verheij, 2005). Our models of abductive reasoning, practical reasoning, and argumentation schemes and diagrams show how motive evidence is distinctively different from other kinds of evidence. A motive is a particularly strong goal or intention that the agent has a strong emotional commitment to implementing. In efforts to

reconstruct the agent's plan, the motive is a special type of goal (or intention) that was the mainspring of the action. It is classified as a goal under the commitment model of practical reasoning, or as an intention under the BDI model of practical reasoning. Here we differentiate, following Carberry (1990 p. 76), between general goals of an agent and immediate goals related to an agent's specific actions that heuristics can identify at a specific segment of a dialogue. A motive is one of these immediate goals that heuristics (especially practical reasoning) linked to a specific action.

On this theory, motive is different from intention, because motive is not an essential element of any charge claim or defense (Leonard 2001, p. 439) in the same way that intention is. As Federal Rule of Evidence 404 shows, evidence of other crimes is admissible if it shows motive, opportunity, intent, preparation, planned, knowledge, identity, or absence of mistake or accident. On our theory, these elements relate to the plan an agent might be thought to have had in carrying out an action. In evaluating their probative value, our theory suggests that the focus, should be on reconstructing the chain of goal-directed reasoning that was the agent's plan, from evidence of what the agent said and did, along with other relevant evidence of the circumstances of the action.

Inferences about character from evidence of past behavior rely on perceptions of a general pattern or habitual mode of action that have an ethical aspect, which the decision maker infers through practical reasoning from an agent's actions and commitments (Walton, 2006).<sup>8</sup> In this respect, it is similar to motive evidence. The key difference, however, is that judging character employs a different bottom-up chain of reasoning. To judge motive we begin with a particular action and use backward abductive reasoning to reconstruct a hypothetical plan derived from the ostensible goals and commitments of the agent. This process employs the same kind of sequence of backward abductive reasoning used to draw an inference about an agent's character. Judgments about motive and those about character are different kinds of conclusions, however. A hypothesis about character necessarily relies on a broad generalization about the agent's values, and about how the agent's past actions have implemented those values. Judgments of character involve a balancing of goals against particular cases in which the agent either followed these goals or failed to do so. Assessment of evidence of a motive is different, because it starts with a specific action and then reasons backwards to a hypothesis about the particular goal that led the agent to decide to take the action, as a result of its practical reasoning. The distinction between general (higher-level) goals and direct goals (linked at a lower level to a specific action in question) is vital.

#### Acknowledgements

Douglas Walton would like to thank the Social Sciences and Humanities Research Council of Canada for a research grant, and Burkhard Schafer would like to acknowledge that his work on this paper was supported by the ESRC grant RES-000-23-0729.

#### References

Atkinson, Katie, Trevor Bench-Capon and Peter McBurney, 'Justifying Practical Reasoning', *Proceedings of the Fourth International Workshop on Computational Models of Natural Argument* (CMNA 2004), ECAI 2004, Valencia, Spain, pp. 87-90.

Atkinson, Katie, and Trevor Bench-Capon, 'PARMENIDES: Facilitating Democratic Debate', *Electronic Government*, ed. R. Traummüller, Lecture Notes in Computer Science (LNCS), 3183. Third International Conference on eGovernment (EGOV 2004), DEXA 2004, Zaragoza, Spain, 2004a.

Atkinson, Katie, and Trevor Bench-Capon, 'Agent Decision Making Using Argumentation About Actions', Technical Report ULCS-05-006, University of Liverpool, Computer Science Department, 2005.

---

<sup>8</sup> Here, for reasons of a lack of space, we can say little more about the logical structure of character evidence in law, but having concentrated on revealing the logical structure of motive evidence, we can point to the essentials of the contrast between it and character evidence (Walton, 2006, pp. 43-46).



Atkinson, Katie, and Trevor Bench-Capon, 'Computational Representation of Practical Argument', *Synthese*, 152, 2006, 157-206.

Audi, Robert, *Practical Reasoning*, London, Routledge, 1989.

Bench-Capon, Trevor, 'Persuasion in Practical Argument Using Value-based Argumentation Frameworks', *Journal of Logic and Computation*, 13, 2003, 429-448.

Bratman, Michael, *Intention, Plans and Practical Reason*, Cambridge, Mass., Harvard University Press, 1987.

Bratman, Michael E., David J. Israel and Martha E. Pollack, 'Plans and Resource-bounded Practical Reasoning', *Computational Intelligence*, 4, 1988, 349-355.

Carberry, Sandra, *Plan Recognition in Natural Language Dialogue*, Cambridge, Mass., The MIT Press, 1990.

Clarke, David S., Jr., *Practical Inferences*, London, Routledge and Kegan Paul, 1985.

Farer, Tom J., 'Restraining the Barbarians: Can International Criminal Law Help?', *Human Rights Quarterly* Volume 22, Number 1, February 2000, pp. 90-117.

Farmer, Lindsay, 'Notable Trials and the Criminal Law in Scotland and England, 1750–1950', in P. Chassaigne and J-P Genet (eds), *Droit et Société en France et en Grande-Bretagne (XIIe-XXe siècles). Fonctions, usages et représentations* Paris, 2003 pp. 149–170.

Federal Rules of Evidence, U. S. Courts, The Federal Judiciary, April 25, 2006, at: <http://www.uscourts.gov/rules/newrules4.html>

Freeman, James B. 'The Appeal to Popularity and Presumption by Common Knowledge', *Fallacies: Classical and Contemporary Readings*, ed. Hans V. Hansen and Robert C. Pinto, University Park, Pa., The Pennsylvania State University Press, 1995, 263-273.

Girle, Rod, David Hitchcock, Peter McBurney and Bart Verhiej, 'Decision Support for Practical Reasoning: A Theoretical and Computational Perspective', *Argumentation Machines: New Frontiers in Argument and Computation*, ed. Chris Reed and Timothy J. Norman, Dordrecht, Kluwer, 2003, 58-84.

Goldman, Alvin, *A Theory of Human Action*, Englewood Cliffs, Prentice-Hall, 1970.

Gordon, Thomas F., 'Introductory Tutorial: Artificial Intelligence Models of Legal Argumentation', ECCO Competence Center, Fraunhofer FOKUS, Berlin, 2003.

Gordon, Thomas F. and Nikos Karacapilidis, 'The Zeno Argumentation Framework', *Proceedings of the Sixth International Conference on Artificial Intelligence and Law*, Melbourne, Australia, 1997, 10-18.

Govier, Trudy, *A Practical Study of Argument*, 3rd ed., Belmont, Wadsworth, 1992.

Hamblin, Charles L., *Fallacies*, London, Methuen, 1970.

Hamblin, Charles L., 'Mathematical Models of Dialogue', *Theoria*, 37, 1971, 130-155.

Horty, John, and Nuel D. Belnap, 'The Deliberative Stit: A Study of Action, Omission, Ability and Obligation', *Journal of Philosophical Logic*, 24, 1995, 583-644.

Hughes, Patrick and Margaret Sanderson, 'Conspiracy Law and Jurisprudence in Canada: Towards an Economic Approach', *Review of Industrial Organization*, Volume 13, 1998

Hume, David, *Commentaries on the Law of Scotland Respecting Crimes* (4th ed. by BELL, B.R.). Edinburgh, Bell and Bradfute, 1844 (Reprinted by the Law Society of Scotland 1986).

Josephson, John R., and Susan G. Josephson, *Abductive Inference: Computation, Philosophy, Technology*, New York, Cambridge University Press, 1994.

Leonard, David P., 'Character and Motive in Evidence Law', *Loyola of Los Angeles Law Review*, 34, 2001, 439-536.

Lodder, Arno R., *Dialaw: On Legal Justification and Dialogical Models of Argumentation*, Dordrecht, Kluwer, 1999.

Paglieri, Fabio and Cristiano Castelfranchi, 'Arguments as Belief Structures', *The Uses of Argument: Proceedings of a Conference at McMaster University 18-21 May, 2005*, ed. David Hitchcock, Hamilton, Ontario, 2005, 356-367.

Park, Roger C., David P. Leonard and Steven H. Goldberg, *Evidence Law*, St. Paul, Minneapolis, West Group, 1998.

Peirce, Charles S., *Reasoning and the Logic of Things*, ed. Kenneth Laine Kettner, Cambridge, Mass., Harvard University Press, 1992.

Pollock, John L., *Cognitive Carpentry*, Cambridge, Mass., The MIT Press, 1995.

Popper, Karl, *The Open Society and its Enemies*, vol 2, London, Routledge and Kegan Paul, 1966.

Prakken, Henry, 'Relating Protocols for Dynamic Dispute with Logics for Defeasible Argumentation', *Synthese*, 127, 2001, 187-219.

Prakken, Henry, 'Modelling Defeasibility in Law: Logic or Procedure?', *Fundamenta Informaticae*, 20, 2001a, 1-20.

Prakken, Henry, Chris Reed and Douglas Walton, 'Dialogues about the Burden of Proof', Proceedings of the Tenth International Conference on Artificial Intelligence and Law, Held June 6-11, 2005 in Bologna, Italy, New York, The Association for Computing Machinery (ACM), 2005, 115-124.

Preyer, Gerhard and Dieter Mans, 'On Contemporary Developments in the Theory of Argumentation', *Protosociology*, 13, 1999, 3-13.

Reed, Chris and Glenn Rowe, *Araucaria, Version 3*, 2005. Available free at <http://www.computing.dundee.ac.uk/staff/creed/araucaria/>.

Reed, Chris and Timothy J. Norman, *Argumentation Machines: New Frontiers in Argument and Computation*, Dordrecht, Kluwer, 2003.

Searle, John R., *Rationality in Action*, Cambridge, Mass., The MIT Press, 2001. Eco, Umberto and Thomas A. Sebeok, *The Sign of Three: Dupin, Holmes, Peirce*, Bloomington, Indiana University Press, 1988.

Segerberg, Krister, 'Towards an Exact Philosophy of Action', *Topoi*, 3, 1984, 75-83.

Verheij, Bart, 'Dialectical Argumentation with Argumentation Schemes: An Approach to Legal Logic', *Artificial Intelligence and Law*, 11, 2003, 167-195.

Verheij, Bart, *Virtual Arguments*, The Hague, Asser Press, 2005.

Walker, Allan Grierson and Norman MacDonald Lockhart Walker, *The Law of Evidence in Scotland*, Edinburgh, T & T Clark, 2000.

Walton, Douglas, *Practical Reasoning: Goal-Driven, Knowledge-Based, Action-Guiding Argumentation*, Savage, Maryland, Rowman & Littlefield, 1990.

Walton, Douglas, *Abductive Reasoning*, Tuscaloosa, University of Alabama Press, 2004.

Walton, Douglas, *Character Evidence: An Abductive Theory*, Argumentation Library Series, Dordrecht, Springer, 2006.

Walton, Douglas and Erik C. W. Krabbe, *Commitment in Dialogue*, Albany, State University of new York Press, 1995.

Walton, Douglas and Fabrizio Macagno, 'Common Knowledge in Legal Reasoning about Evidence', *International Commentary on Evidence*, 3, 2005, 1-42.

Wigmore, John H., *The Principles of Judicial Proof*, 2nd ed., Boston, Little, Brown and Company, 1931.

Wooldridge, Michael, *Reasoning about Rational Agents*, Cambridge, Mass., The MIT Press, 2000.

Wooldridge, Michael, *Introduction to MultiAgent Systems*, Chichester, Wiley, 2002.

Zizek, Slavoj, The Detective and the Analyst, *Literature and Psychology* 36(1990): 27-46.