Formalising Argumentative Story-based Analysis of Evidence

F.J. Bex
Centre for Law & ICT
University of Groningen
the Netherlands
f.j.bex *at* rug.nl

H. Prakken
Centre for Law and ICT,
University of Groningen
Department of Information and Computing Sciences,
Utrecht University

B. Verheij
Department of Artificial Intelligence
University of Groningen
the Netherlands

ABSTRACT
In the present paper, we provide a formalised version of a merged argumentative and story-based approach towards the analysis of evidence. As an application, we are able to show how our approach sheds new light on inference to the best explanation with case evidence. More specifically, it will be clarified how the events in a case story that are considered to be proven abductively explain the otherwise unproven events of the case story. We compare our approach with existing AI work on modelling legal reasoning with evidence.

Keywords
Argumentation, Abductive reasoning, Inference to the best explanation, Legal evidence, Stories

1. INTRODUCTION
We are developing a conceptual and formal framework for the analysis of reasoning with evidence about the facts of a legal case. The main reason for this is to provide the formal foundations of sense-making software for crime investigation. Such a sense-making system does not contain a knowledge-base, so it does not need domain knowledge. An advantage of this is that sense-making systems are not subject to the knowledge acquisition bottleneck; because reasoning with evidence encompasses a great number of real-world problems, a knowledge based system would need a huge amount of domain knowledge. A sense-making system does not reason itself, but instead helps the user make sense out of a case by structuring it. In the case of a crime investigation it can, for example, structure a case into a timeline and link the different events to police records or photographs. Sense-making systems also exist as more general reasoning aids (see [24] for an overview). The work presented in this paper is part of a larger project in which such a sense-making system is developed for crime investigation applications. Van den Braak and Vreeswijk [4] have developed a prototype system; the present paper aims to investigate the theoretical basis of such systems.

In our formal framework for the analysis of reasoning with evidence we attempt to merge two main approaches of analysing reasoning with evidence: argumentative analysis and story-based analysis.

The former goes at least back to Wigmore's work in the early twentieth century [26] and has been elaborated upon by researchers of the New Evidence Scholarship, such as Anderson, Schum, Tillers and Twining [1], [20]. Recently, two tools have been added to this style of work: argumentation schemes as used in the field of (informal) argumentation theory [25], and argumentation logics, especially in the style as developed in the field of artificial intelligence and law [16]. From the start of applying these tools to evidence in the law, there have been attempts to combine the two tools [2], [23].

The story-based perspective has been taken up by legal psychologists such as Pennington and Hastie [12], which has been followed up in the anchored narratives approach by Crombag, Van Koppen and Wagenaar [5]. Thagard [18] has applied his connectionist model of abductive inference to the best explanation to legal cases, claiming that it provides a computational account of the story approach to legal evidence. Keppens and Schafer [9] have applied model-based reasoning with scenarios to automated support for the investigation of murder cases.

The anchored narratives approach [5] is a starting point of our research. Interestingly, although Crombag, Van Koppen and Wagenaar focus on the story-based perspective in their choice both of wording and of research background, several of their central claims have a more argumentative than story-based flavour. Especially the role of generalisations (or anchors), exceptions to these generalisations and of the dynamics of developing and refining an analysis of the evidence in a case are characteristic for the argumentative slant of the approach by Crombag, Van Koppen and Wagenaar. As a result, the anchored narratives approach can be regarded as a beginning of a merged argumentative and story-based approach. However, as we already noted in [3], the details of what part generalisations play in a story are left untreated. Also, the exact interplay between the evidence, a story and the generalisations connecting these two is not elaborated upon. Accordingly, another aim of this paper is to provide a thorough analytic account of the elements of a story and how a story relates to generalisations and evidence.

The outlines of this merged approach were discussed in our previous paper [3], where we proposed a combination of two techniques from AI, namely inference to the best explanation and argumentation theory. In this paper, we want to extend and

ICAIL '07, June 4-8, 2007, Palo Alto, CA USA.

fully formalise this combined theory of evidential reasoning. We do this by proposing a theory that combines abductive explanations with evidential arguments. To our knowledge, these two formal theories have so far not been integrated in any previous work.

The rest of this paper is organised as follows: in section 2 we discuss the typical aspects of evidential reasoning in crime investigation contexts as far as relevant for this paper. Section 3 informally discusses the two current approaches on which we base our combined theory and argues why these two approaches should be combined. In section 4 we propose our combined, formal theory; section 5 contains an extended example of argumentative, story-based analysis of the evidence using our formal theory. Finally, in section 6 we discuss related research conclude with a discussion and ideas for future research.

2. EVIDENTIAL REASONING

In this section we aim to give an overview of the typical aspects and elements of reasoning with evidence. For this overview, we draw on our earlier work [3], which was in turn based on the work on Anchored Narratives [5]. Note that here we take a static viewpoint: we assume that once the analysis starts, all the available evidence and stories are at hand. In this paper, we do not concern ourselves with dynamic aspects of evidential reasoning like, for example, the search for new evidence and the updating of stories in light of new evidence. We will clarify this overview by means of an example, which we will also use in section 5.

The example concerns a simplified version of the Rijkbloem case ([5], pp. 78). Danny Rijkbloem is a 23-year old man from Surinam. He has a considerable list of sentences (theft, robbery) starting when he was 15 years old. Nicole Lammers is a 20-year old baker's daughter who had a relationship with Rijkbloem and lived together with him. At some point Nicole decided, under pressure of her parents that it is best to break up with Rijkbloem and she leaves him. A few days after the break-up, Nicole and her parents went to Rijkbloem's house to pick up some of Nicole's stuff, and a fight developed between Rijkbloem and Nicole's father. From this point onwards, the two women (Nicole and her mother) and Rijkbloem tell a different story. According to the two women, Rijkbloem pulled out a gun with which he shot father Lammers through the head from a distance of about 2 meters. Rijkbloem, however, gives a different account of the events. He said that during the fight, Mrs. Lammers pulled a gun out of her purse and threatened to shoot Rijkbloem with it. Rijkbloem pushed the hand holding the gun away and in the struggle the gun went off and the bullet hit father Lammers in his head.

In both crime investigation and legal decision making, the investigators or decision makers are faced with a number of *stories*. Such a story is usually a set of *states and events*[1] structured in a chronological order. Sometimes this story presents itself through the victim's testimony. In other cases the investigators will have to construct a possible story by means of the physical evidence that is presented to them. In the Rijkbloem case, the police was presented with two stories: The account of the mother and daughter that Rijkbloem shot the father and Rijkbloem's account, that it was the mother who accidentally shot the father. Stories are often used to "fill the gaps": not every part of a story follows directly from evidence, but the story as a whole is plausible, so we can more or less safely assume that the events in the story that do not follow from evidence also happened.

The second important element of evidential reasoning is *evidence*. Evidence can be physical evidence, like a knife or a body, or testimonial evidence (this classification of types of evidence is taken from [26]). In the Rijkbloem case, the police had as physical evidence the father's body. Note that a forensics report on the father's cause of death is considered testimonial evidence. Other testimonial evidence in the Rijkbloem case was the mother and daughter's testimony and Rijkbloem's testimony.

Once the stories and the evidence are gathered, the analysis of the case can start. The first thing that has to be done is deciding which events are to be explained, the *explananda*. The main explananda are usually the observations that give rise to the investigation or the legal case in the first place, for example, "how did this man die?" or "who robbed the bank?" Often, new explananda appear during the investigation or treatment of a case, for example, "why was this man acting suspiciously near the crime scene?" Explananda usually follow from the evidence in a case. In the Rijkbloem example the most important explanandum is the father's dead body; what caused the father to die?

Often, there are only a few stories or scenarios that explain the most important explananda. In the Rijkbloem example, there are two stories that explain the father's death. For each of these stories, the investigator or decision maker should look which of the events in the separate stories are supported by evidence and which events are not supported by evidence. So, for example, the father's body supports the fact that he is dead and the forensics report supports that the father was shot from a distance of 50 centimetres. Rijkbloem's testimony supports, of course, most of the events in his version of the story while the women's testimonies support their own story. Other testimonies by, for example, acquaintances of the Lammers family support the facts that Nicole and Rijkbloem had a relationship.

When the different stories and their supporting evidence are identified, the *generalisations* must be analysed, as the generalisations are the "anchors" of the different stories. There are (at least) two distinct types of generalisations, namely *causal* and *evidential*.

The causal generalisations are the glue of a story; they give a story its coherence. For example, the fact that "Rijkbloem shot the father so the father died" is a coherent part of the women's story is because we believe the causal generalisation "if *x* shoots *y*, this can cause *y* to die". The evidential generalisations concern the connection between the elements of a story and the sources of evidence. In order for a forensics report to support the (story) event that the father was shot from a distance of 50 centimetres, we have to believe that "usually, a forensics report describing event *e* is evidence for the fact that event *e* really happened".

[1] It can be argued that an occurrence of an event is also a state. For a discussion of this issue, see [8]. In this paper, a *state* is a state of affairs in the world, and an *event* is something that causes a change in the obtaining states of affairs.

When all these elements have been thoroughly analysed, it is up to the investigator(s) or decision maker(s) to decide what story to believe. Ideally, this will be the story that is best supported by evidence and contains no dubious generalisations.

3. ARGUMENTATIVE AND STORY-BASED ANALYSIS

In this section we first informally discuss two dominant approaches to reasoning with legal evidence in AI and Law, namely argumentation and inference to the best explanation or abductive-causal reasoning. Then, in section 3.2, we will discuss some of the advantages and disadvantages of the two approaches and argue that a combined approach best captures the aspects of evidential reasoning as discussed in section 2.

3.1 Current approaches to reasoning with legal evidence

In logics for argumentation (see [16] for an overview), the rules of classical logic are augmented with rules for defeasible inference. Associated with a defeasible inference is an underlying generalisation that acts as a *warrant* [21] (cf. anchor [5], rule [7] and scheme [25]). Arguments can be constructed by chaining applications of inferences and thus one ends up with a tree of arguments akin to a Wigmore graph [26], [1]. Each tree has as its premises the pieces of evidence and the generalisations and as its ultimate conclusion one of the explananda (in a Wigmore graph, the generalisations are contained in the links). In Figure 1, an example of an argument tree is given, where the "E" nodes are the pieces of evidence, the "G" nodes are evidential generalisations and the arrows are inferences. The other nodes are (intermediate) conclusions.

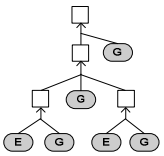

Figure 1: an argument tree

Notice that the inferences in such an argument are all of an *evidential* nature: a piece of evidence *e* and the evidential generalisation "*e* is evidence for *p*" allows us to infer *p*. For example, the evidence that "Nicole saw that Rijkbloem shot her father" and the generalisation that "if a witness sees an event, this is evidence for the occurrence of that event" allow us to infer that "Rijkbloem shot the father".

Arguments can also be *attacked*. They can be *rebutted* with an argument for the opposite conclusion and they can be *undercut* with an argument for why an inference is not allowed (usually because a generalisation does not apply in the given circumstances). In the example, an argument for "Rijkbloem did not shoot the father" rebuts "Rijkbloem shot the father" and an argument for "Nicole wants to protect her mother" undercuts the generalisation "if a witness testifies that an event happened, this is evidence for the occurrence of that event" as applied to Nicole's testimony. When these attack-relations are known, it can be determined which arguments "win", which arguments "lose" and which are "undecided".

In the argumentation approach, different arguments like the one shown in Figure 1 are built, supporting the propositions that have to be proven, and other arguments are also built to attack these arguments. In the example, the prosecution has to prove that Rijkbloem killed the father. In order to do this, arguments will have to be constructed for, for example, "Rijkbloem and the father were in the same house", "Rijkbloem had a gun which he used to shoot the father". The defense will try to attack these arguments by saying, for example, that it was not Rijkbloem but the mother who shot or by arguing that the testimonies of the two women cannot be trusted.

In story- or explanation-based approaches, a story (also referred to as a scenario) is modelled as a causal network. Figure 2 depicts a small part of the prosecution's story in the Rijkbloem case.

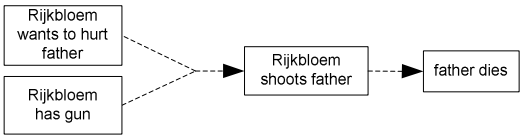

Figure 2: a small causal network

The nodes in this graph are events in the story and the links represent causal relations between the events: Rijkbloem shooting the father causes the father to die.

Such a causal network can be used as a *causal theory*, in which observations are *explained* by hypothesised events or states through abductive inference to the best explanation (IBE). The basic idea of abductive inference (for an overview see [10]) is that if we have a general rule *cause* → *effect* and we observe *effect*, we are allowed to infer *cause* as a possible explanation of the effect. This cause which is used to explain the effect can be a single state or event, but it can also be a sequence of events, a story. In the example, the observation that has to be explained is "father dies". This can be explained by the causal theory consisting of the generalisations "If person X wants to hurt person Y and person X has a gun, this will (usually/sometimes) cause person X to shoot person Y" and "Person X shooting person Y (usually/sometimes) causes person Y to die" together with the assumptions that "Rijkbloem wants to hurt father" and "Rijkbloem has a gun". The theory together with these initial causes is then a hypothesis which explains the observation "father dies".

Taken by itself the abductive scheme is nothing but the fallacy of affirming the consequent. However, in a setting where alternative abductive explanations are generated and compared, it can still be rational to accept an explanation if no better other explanation is available. Clearly, such reasoning is defeasible, since additional facts might give rise to new explanations. In sum, the idea of abductive inference to the best explanation is that there are different explanations which have to be compared.

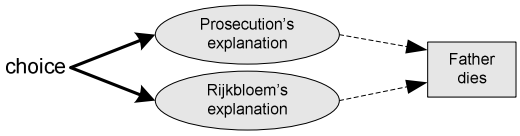

Figure 3: two explanations

In the Rijkbloem case, we have two different stories: the prosecution's story that it was Rijkbloem who shot the father, and Rijkbloem's story that it was the mother who (accidentally) shot the father. This is visualised in Figure 3.

In the literature there are different ways of comparing the explanations. The simplest way is choosing a subset-minimal explanation that explains all the observations. Thagard [18] computes the activation of the individual propositions: (causal) coherence between two units functions as an excitatory link between the units. Acceptance or rejection of an explanation is represented by the degree of activation of the individual propositions in the explanation.

3.2 Why combine the two approaches?

In this section we argue why we think a combination of the two approaches discussed in the previous sections best captures the concepts expressed in section 2. We will do this by briefly discussing the advantages and disadvantages of the two approaches when applied to reasoning with evidence.

Advantages of the argument-based approach are that it is based on simple principles of reasoning and that it is transparent: the generalisations and the evidence are mentioned explicitly in the arguments and the reasoning from the evidence to the conclusion is easy to follow. A disadvantage of the argument-based approach is that it does not provide a complete overview of the case, as the original stories about "what happened" are cut into pieces to become conclusions of different arguments and counter-arguments. It may be possible to reason abductively within a purely argument-based approach by using the abductive scheme (if *cause* → *effect* and *effect* then usually *cause*) as a warrant for making the inference. However, this kind of abductive reasoning is evidential in nature (see last paragraph of this section), so by using the purely argument-based approach one loses the causal information and the causal reasoning inherent to explanations.

An important advantage of the IBE approach is that it perhaps is closest to how legal decision makers actually think about a case. Experiments by Pennington and Hastie [12] suggest that when making a decision, people often compare the different stories from which the explananda follow instead of constructing complex arguments for and against the explananda. Choosing the best explanation is closer to this 'holistic' approach to reasoning with events, where an event is evaluated by looking at how well it fits a coherent story (see [22], pp. 241 for a short discussion on holistic and atomistic reasoning with evidence).There are also indications that fact investigators work with causal structures and timelines [15].

However, the IBE approach is arguably less natural in its treatment of sources of evidence. In IBE approaches, pieces of evidence are modelled as effects. So a witness testimony, for example, must be abductively explained by its possible causes, which are other states or events. An obvious explanation is that the testimony was caused by the fact that the event described in the testimony happened. However, there may be other explanations, such as that the witness tried to protect a suspect or that the witness's senses or memory are at fault. In a purely abductive approach all such explanations always have to be considered. However, in reality the 'truthful' explanation is considered the normal one and alternative explanations are regarded as exceptions to the generalisation that witnesses usually speak the truth, and are investigated only if there is evidence for such exceptions. In [3] we argued that this is better captured by modelling reasoning with sources of evidence as argumentation with evidential default generalisations. The witness generalisation would then be modelled as "If a witness says that "P" then usually P is true", and the undercutting exceptions to this generalisation are, for example, "the witness tries to protect a suspect" or "the witness' memory is at fault".

Another problem of the IBE approach as it is usually modelled is that it is impossible to reason *about* the causal generalisations in the causal theory. When IBE is used in a diagnostics tool, the causal theory is determined by experts in the relevant field, for example, medicine in the case of a medical diagnostics system. In that case, it is reasonable to assume a causal theory based on expert consensus and thus the explanations that are provided for a given explanandum. In the case of reasoning with evidence, however, we cannot assume that the causal theory provided is correct or agreed upon; in fact, in an argumentative story-based proof analysis the analysis and perhaps change of the causal generalisations in the stories is a central point of focus.

A final issue of the IBE approach is the comparison of explanations. As Thagard and Shelley [19] noted, the simplest (in terms of subset minimality) and most complete (in terms of explaining the most observations) explanation is not always the best. Particularly in reasoning with evidence, a complex explanation that explains only a few important pieces of evidence can be better than a simple explanation that explains many less important pieces of evidence. A more fine-grained approach to comparing explanations is needed for reasoning with evidence. The evidential arguments will have a big influence here; these arguments allow one to decide which evidence to believe and which evidence is most important, and thus which evidence needs to be explained.

In our opinion, an approach that combines the argumentation approach and the IBE approach best models the concepts of argumentative, story based analysis of evidence as discussed in section 2. In this combined approach, the stories are represented as networks of causally linked events. It should be noted that we use a naïve interpretation of causality; in our examples, sometimes a link does not represent a much stronger relation than temporal precedence. Because we also allow arguments that causal generalisations in the model are attacked, this interpretation of causality does not pose a real problem; it is up to the investigator or the decision maker to change the causal theory into a correct theory. This approach to causality also allows for the networks to remain relatively simple, so that they keep providing a good overview of the case.

In our approach, the *sources of evidence* are connected to events by *evidential arguments*, with pieces of evidence as premises and events as conclusions, and *evidential links*, representing inferences, that connect the evidence to the event. This gives the evidence a clear and separate place in the theory, something which was lacking in other, the Anchored Narratives approach [5]. It also solves the problems the IBE approach has with its treatment of sources of evidence as discussed in section 3.2. A third advantage of adding arguments to reasoning with

explanations is that the validity of the causal generalisations in a theory can be argued about.

Before we continue with the formal theory, an important remark should be made. When we talk about evidential links, we mean only the links in the argument that connect the source of evidence to the event. Other authors, notably Pearl [11] and Poole [14], speak of evidential links or evidential reasoning in a broader sense. According to them, causal reasoning is reasoning from cause to effect, where a certain cause *predicts* an effect, and evidential reasoning is reasoning from effect to cause, where a cause *explains* an effect. Evidential reasoning can be done through abduction, viz. "*cause* → *effect* and *effect* so cause" or deduction ("modus ponens style reasoning"), viz. "*effect* → *cause* and *effect* so *cause*". In this paper, we use both kinds of evidential reasoning: the causal networks representing the stories *abductively explain* the pieces of evidence; and sources of evidence *deductively explain* why we should believe that certain events happened. Of course, here abductive and deductive reasoning are both considered defeasible. However, when we refer to evidential reasoning, we mean the deductive evidential inferences from the sources of evidence to the events. Our approach respects Pearl's c-e rules [11], that is, in our approach it is not allowed to apply an evidential generalisation to a proposition that has been inferred by application of a causal generalisation.

4. A FORMAL THEORY

In this section we will present our combined formal theory. Section 4.1 discusses the argumentation logic and section 4.2 discusses how the causal theories are used to explain the explananda. In section 4.3 we discuss how different explanations can be compared by taking into account the evidential arguments.

Combining the causal theories and the evidential arguments, we can define a framework for argumentative, story based evidence analysis. This framework is a combination of an explanation framework and an argumentation framework. An *explanation framework* is a tuple $A_C = (T, H, O, F)$. Here, T is the *causal theory* which contains all the causal generalisations from the different stories. H is a set of *explanantia*, propositions with which we want to explain the explananda. The set of *observations* O contains all the propositions that are supported by some evidence (through an evidential argument). $F \subseteq O$ is the set of *explananda*, which are ideally explained. The idea is that the explanantia H together with a part of the causal theory T explain the explananda F. An *argumentation framework* is a pair $A_E = (G_E, E)$, where G_E is a set of *evidential generalisations* and E is a set of *evidence*. Combining these two frameworks, our framework for evidential reasoning is a pair $ER = (A_C, A_E)$ such that $O \subseteq c(A_E)$, where $c(A_E)$ are the conclusions on the basis of A_E.

The generalisations in G_C and G_E are formalised with a special conditional connective ⇒ which only satisfies the modus ponens inference rule.

$\mathbf{g_i}(x)$: $p_1 \wedge ... \wedge p_n \Rightarrow q$

Here $\mathbf{g_i}$ is the name of the generalisation, x is a tuple of variables denoting the terms in the generalisation and $p_1 \ldots p_n$ and q are predicate-logical literals. The type of generalisation is indicated with a subscript: $\Rightarrow_E$ denotes an evidential generalisation and $\Rightarrow_C$ denotes a causal generalisation. In the following subsections we will elaborate on our approach.

4.1 A Logic for Argumentation

Since reasoning about evidence is defeasible, the logic $\mathscr{L}$ of our system must be a nonmonotonic logic. Following our earlier work in [2], we augment the inference rules of classical logic with defeasible inference rules (cf. Pollock [13]). In this paper, we only need one defeasible inference rule, namely the just-mentioned modus ponens rule for the conditional operator ⇒. The application of this inference rule implicitly assumes that there is no exception to the generalisation to which the inference rule is applied and that this generalisation is valid. Such an application can thus be attacked by arguing that there is an exception or that the generalisation is not valid. An exception to an evidential generalisation $\mathbf{ge_i}$ is modelled as $exc(\mathbf{ge_i}(x))$, where $exc(\mathbf{ge_i}(x))$ stands for "there is an exception to generalisation $\mathbf{ge_i}(x)$", and $\neg valid(\mathbf{ge_i}(x))$ means that generalisation $\mathbf{ge_i}(x)$ is not valid. This modeling of exceptions is related to Hage's [7] approach to rule exceptions and validity.

In his work, Pollock defines specific inference rules relevant for evidential reasoning, such as perception, memory and temporal persistence. Moreover, in [2] we formulated reasons for expert and witness testimonies. In this paper, however, we will model these reasons using the general inference rule, but we still retain the option of modelling them as separate inference rules for future work.

The evidence E and the evidential generalisations from G_E allow us to build evidential arguments by taking the evidence as premises and chaining applications of defeasible modus ponens into tree-structured arguments. Such an *evidential argument* is a finite sequence of lines of argument, where a line is either a proposition from E or the result of an application of the defeasible modus ponens to one or more previous lines. Take, for example, the following evidential generalisations

$\mathbf{ge_1}(W, P)$: W testifies that "P" $\Rightarrow_E$ P is true
$\mathbf{ge_2}(W, P)$: W saw that P happened $\Rightarrow_E$ P happened.

These two evidential generalisations can be used to construct argument AR_1 for the conclusion that Rijkbloem shot the father:

1. $\mathbf{e_1}$: Nicole testifies: "I saw Rijkbloem shoot my father"
2. $\mathbf{ge_1}$: W testifies that "P" $\Rightarrow_E$ P is true
3. Nicole saw Rijkbloem shoot the father *(1, 2, def. MP)*
4. $\mathbf{ge_2}$: W saw that P happened $\Rightarrow_E$ P happened
5. Rijkbloem shot the father *(3, 4, def. MP)*

Note that we will often shorten the name of a generalisation $\mathbf{ge_i}(x)$ to $\mathbf{ge_i}$. An argument can defeat another argument by rebutting or undercutting the other argument. Two arguments *rebut* each other if they have the opposite conclusion. An argument AR_i *undercuts* another argument AR_j if there is a line $exc(\mathbf{ge_i}(x))$ or $\neg valid(\mathbf{ge_i}(x))$ in argument AR_i and a line $Q(x)$ in argument AR_j, which is obtained from some previous lines in AR_j by the application of defeasible modus ponens to $\mathbf{ge_i}(x)$: $P(x) \Rightarrow Q(x)$.

So an argument A_1 defeats another argument A_2 if A_1 rebuts one of A_2's (sub)conclusions or if A_1 undercuts an instance of a generalisation in A_2. As an example of rebuttal, take the following argument AR_2:

1. $\mathbf{e_2}$: Rijkbloem testifies: "I did not shoot the father"

2. Rijkbloem did not shoot the father *(1, **ge**$_1$, def. MP)*

AR_2's conclusion that "Rijkbloem did not shoot the father" rebuts AR_1's conclusion that "Rijkbloem shot the father" (and vice versa). Note that in the remainder of this paper, we will not reproduce the evidential generalisations in a line of argument, but simply refer to them by their name. To give an example of an undercutting argument, we first need a generalisation that points to an exception:

ge$_3$ (W): ¬(W is trustworthy) $\Rightarrow_E$ exc(**ge**$_1$(W, y))

This generalisation expresses that if a witness W is not trustworthy, generalisation **ge**$_1$ is not applicable anymore (insofar as **ge**$_1$ is about W). With new evidence an undercutting argument AR_3 can be built:

1. **e**$_3$: Rijkbloem testifies that "Nicole is not trustworthy"
2. exc(**ge**$_1$(Nicole, y)) *(1, **ge**$_3$, def. MP)*

This argument undercuts AR_1.

When we have a collection of arguments and their binary defeat relations, the dialectical status of the arguments must be defined. In this paper, we assume that our semantics instantiates Dung's [6] grounded semantics. In these semantics, arguments can be either *justified*, which means that they are not attacked by other justified arguments that are stronger, or *overruled*, which means that they are attacked by one or more other stronger arguments that are justified, or *defensible*, which means that they are neither justified nor overruled. Note that in the present paper, we will not discuss the relative strength between arguments.

4.2 Explanations

The IBE part of our theory is mostly based on standard accounts of inference to the best explanation and abductive reasoning [10]. The causal theory T is the collection of all causal generalisations that appear in the different stories. An effect can be a cause of some other event and thus causal generalisations can be chained. As an example, take the following causal generalisations

gc$_1$(x, y) : x shoots y $\Rightarrow_C$ y is hit in the head
gc$_2$ (y) : y is hit in the head $\Rightarrow_C$ y dies

These two generalisations can be chained, and in natural language this would read as: "if a person (x) shoots another person (y), this may cause the person (y) to be hit in the head and if someone is hit in the head, this may cause that person to die".

The set of observations O contains the propositions that follow from an evidential argument. For example, if there is an argument for the proposition "Rijkbloem shoots father" and **gc**$_1$ is part of T, then "Rijkbloem shoots father" will be in O.

In standard accounts of IBE the observations are usually given and they all have to be explained. One of the main ideas in our formal theory is that the set of observations is not simply given but is determined by the conclusions of evidential arguments. Furthermore, not all observations have to be explained. Instead only the observations which are in the set of explananda F have to be explained. In the Rijkbloem example, the event that "Rijkbloem and Nicole broke up" is supported by an argument but it is not a part of F. This is because we are not interested in an explanation for the fact that they broke up; we want to know who shot the father. So the set of explananda F is essentially chosen by the investigator or decision maker (with the restriction that they have to be supported by an evidential argument). It is possible to divide the explananda into propositions which *have to be explained* and *additional evidence* which can be explained. Propositions in F which are supported by a *justified* argument have to be explained and propositions which are supported by a *defensible* argument are ideally explained, but this is not necessary. Note that there will always be at least one proposition in F that is supported by a justified argument; otherwise there is no point in looking for an explanation.

The explanantia H are also chosen by the investigator or decision maker. Now the precise definition of explanations can be given.

Definition (explanations)

$S_i = H_i \cup T_i$, where $H_i \subseteq H$ and $T_i \subseteq T$, is a proper *explanation* of the explananda F iff

1. $\forall f$: If $f \in F$ and f is the conclusion of a *justified* (sub)argument that follows from A_E, then:
   - $H_i \cup T_i \vdash f$; and
   - $H_i \cup T_i$ is consistent.
2. There is no justified argument for the conclusion ¬valid(**gc**$_i$) or exc(**gc**$_i$), where **gc**$_i \in T_i$.

Here $\vdash$ stands for logical implication according to the set of all deductive inference rules extended with modus ponens for $\Rightarrow$. So the hypothesis H_i (which contains some or all of the explanantia) together with the (sub)theory T_i should explain all explananda which are supported by a justified argument. Note that there should be no justified argument for the conclusions that one of the generalisations in the (sub)explanation is invalid or that there is an exception to it. The idea behind this second point in the definition is that we should not be able to explain propositions using a theory that contains invalid generalisations or generalisations to which an exception has been provided; this part of the definition allows us to reason about the generalisations in the explanation. However, care must be taken when dealing with arguments with an exception to a generalisation or the invalidity of a generalisation as their conclusion. One of the reasons for this is that at present, generalisations themselves do not have strengths, so they are always automatically "defeated" by an argument for the opposite conclusion. If the causal generalisations have strengths, it can be imagined that, for example, a "strong" causal rule (e.g. x kicks rubber ball $\Rightarrow_C$ ball moves; most rubber balls move when you kick them) is less easily declared not valid than a "weak" causal generalisation (e.g. x has fight with y $\Rightarrow_C$ x kills y; not everybody who has a fight with another person kills the other person). Because of this subtlety, we allow only justified arguments to "block" an explanation.

As a small example, take T = {**gc**$_1$, **gc**$_2$} (see the beginning of this section), H = {Rijkbloem shot father, mother shot father}, F = {father died} and suppose that there is a justified argument for father died. Then there are two explanations, namely S_r = {Rijkbloem shot father} $\cup$ T and S_m = {mother shot father} $\cup$ T. The explanandum follows from both explanations and both explanations are consistent.

4.3 Comparing explanations

Our combination of two reasoning models makes a more refined assessment of stories possible than in the two individual approaches. Not all of the observations have to be explained: the most interesting observations can be chosen as explananda. Also, in our formal proposal causal links are not just given but their validity can be argued about and exceptions can be given to causal generalisations. In this section, we will further discuss how different explanations can be compared. The criteria discussed in this section are meant to be guidelines for analysing explanations. We do not claim that the decision about which explanation is the best is easily made using these criteria. Other factors, like the quality of the generalisations in an explanation, also play a role.

From the definition of explanations in section 4.2 it follows that an explanation has to explain all the explananda which follow from a justified argument, but it does not have to explain all the additional evidence, that is, observations which follow from defensible arguments. The rationale behind this is that if the argument is justified, we are more certain the event really happened, and we are less sure that events which follow from defensible arguments happened.

However, the explananda that follow from defensible arguments are not overruled and thus can provide us with a criterion for the comparison of explanations: an explanation is preferred (over other explanations) if it explains more propositions in F which follow from *defensible* evidential arguments. As an example, take T = {**gc**$_1$, **gc**$_2$, **gc**$_3$}, where **gc**$_3$: x shoots $y \Rightarrow_C x$ has gunpowder on his hands. Furthermore, suppose that H = {Rijkbloem shot father, mother shot father}, F = {father died, Rijkbloem has gunpowder on his hands}, the argument for father died is justified and the argument for Rijkbloem has gunpowder on his hands is defensible. Now there are still two explanations S_r and S_m (see end of section 4.2), but S_r also explains Rijkbloem has gunpowder on his hands, making it a better explanation than S_m.

In addition to there being arguments for the explananda, there can, of course, also be evidential arguments for other events in an explanation. This brings us to the second criterion: the more propositions in an explanation are supported by a non-overruled argument, the better an explanation is. Normally, the propositions in an explanation are inferred abductively from an observation: if we have the causal generalisations a $\Rightarrow_C$ b, b $\Rightarrow_C$ c and an observation c that we want to explain, we can abductively infer first b and then a. Even though abduction is not logically valid, the idea is that we are safe to assume that a together with the generalisations is an explanation of c and that thus a and b will be the case. This is the "gap filling" function of stories (see section 2) in effect: a and b do not follow from evidence, but the explanation as a whole is plausible, so we can more or less safely assume that a and b are the case. If, however, a and b from the above example are also supported by an evidential argument this is, of course, even better; they can not only be inferred abductively but also follow from evidence.

It is also possible that the conclusion of an evidential argument *contradicts* an event in an explanation. This brings us to the third criterion for comparing explanations: an explanation is preferred if it contradicts fewer conclusions of non-overruled arguments. With this criterion it does not matter whether the argument which conclusion contradicts the explanation is justified or defensible. One could argue that an explanation that contradicts a defensible conclusion is better than an explanation that contradicts a justified conclusion; however, in this paper we choose not to make this distinction. As an example, take the situation as laid out at the end of section 4.2 together with a non-overruled argument for Rijkbloem did not shoot the father (AR_2 on page 5). This time, explanation S_m is better, because S_r contradicts the conclusion of argument AR_2.

5. EXAMPLE

In this section, we present an example of how different elements of a case can be analysed using our merged theory of argumentative, story-based analysis of evidence. It should be noted that the examples in this section present different time points in an ongoing analysis of one case. As said, we have not yet defined a dynamic framework for argumentative, story-based analysis.

We start our analysis with the two stories told by Rijkbloem and Nicole and her mother (see page 2).

T = { **gc**$_1$: fight between x and $y \wedge x$ has gun $\Rightarrow_C x$ shoots y,
gc$_2$: x shoots $y \Rightarrow_C y$ is hit,
gc$_3$: y is hit $\Rightarrow_C y$ dies,
gc$_4$: fight between x and $y \wedge z$ has gun $\Rightarrow_C z$ aims gun at x,
gc$_5$: x aims gun at $y \Rightarrow_C y$ tries to push gun away,
gc$_6$: x tries to push y's gun away $\Rightarrow_C y$'s gun accidentally goes off,
gc$_7$: gun accidentally goes off $\Rightarrow_C x$ is hit}

F = {father dies}

H = {fight starts between Rijkbloem and father, Rijkbloem has gun, mother has gun}

G = {**ge**$_1$: W testifies that "P" $\Rightarrow_E P$ is true
ge$_2$: W_1 and W_2 testify that "P" $\Rightarrow_E P$ is true
ge$_3$: police report that "P" $\Rightarrow_E P$ is true}

E = {**e**$_R$: Rijkbloem testifies that:
- fight starts between him and father
- mother aims gun at him
- he tries to push gun away
- gun accidentally goes off
- father is hit in the head

e$_W$: Nicole and her mother testify that:
- fight starts between Rijkbloem and father
- Rijkbloem shoots father
- father is hit in the head

e$_P$: police report that "Mr. Lammers died"}

O = {fight starts between Rijkbloem and father, mother aims gun at Rijkbloem, Rijkbloem tries to push gun away, gun accidentally goes off, Rijkbloem shoott father, father is hit in the head, father dies}

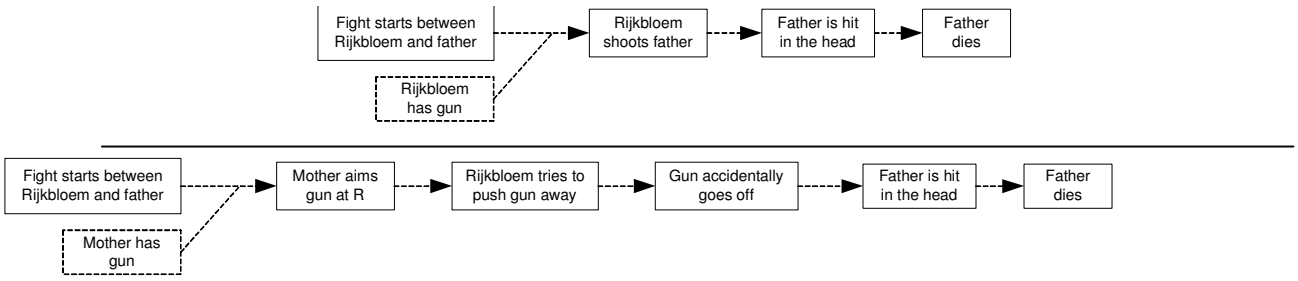

Figure 4: the two explanations S_m and S_r

With G and E, evidential arguments can be constructed for the propositions in O. All these arguments are justified because none of them is attacked by another argument. With this data, two explanations can be made. The first explanation is the women's story, that Rijkbloem was the one who shot the father: $S_r = T_r \cup H_r$, where T_r = {**gc₁**, **gc₂** **gc₃**} and H_r = {fight starts between Rijkbloem and father, Rijkbloem has gun}. The second explanation is Rijkbloem's story that the mother's gun went off: $S_m = T_m \cup H_m$ where T_m= {**gc₄**, **gc₅**, **gc₆**, **gc₇**} and H_m = {fight starts between Rijkbloem and father, mother has gun}. The explanations have been represented as graphs in Figure 4. The two propositions that are not supported by evidence are represented as boxes with dotted lines.

It is hard to make a decision which of these two explanations is better if we follow the criteria from section 4.3; both explanations are supported by evidence, and in both explanations there is an important proposition which is not supported by evidence, namely "Rijkbloem/mother has a gun".

However, the women's story seems more plausible: a caring father tries to protect his daughter from a criminal and pays with his life. Fathers want to protect their daughters at all costs. Likewise, when a problem arises, criminals such as Rijkbloem resort to violence. Rijkbloem can have made up his strange story to protect himself; women do not normally carry guns around, and the chance that the bullet hit the father after a "struggle" between Rijkbloem and the mother is low. This shows how a "good" story can win over a "bad" story. Below we will formalise this idea and also show that the most plausible story is not always better supported by evidence.

First we want to argue that Rijkbloem made up his story in order to protect himself:

1. Rijkbloem is a suspect
2. *x* is a suspect $\Rightarrow_E$ *x* wants to protect himself
3. Rijkbloem wants to protect himself (*1, 2, def. MP)*
4. *x* wants to protect himself $\Rightarrow_E$ exc(**ge₁**(*x, p*))
5. exc(**ge₁**(Rijkbloem, *p*)) *(3, 4, def. MP)*

This argument undercuts every argument that has e_R as its premise; in other words, Rijkbloem's story is not believed. This will make S_r (the story that Rijkbloem shot the father) the preferred explanation, because all arguments which support S_m are undercut. It is in this way that the interactions between the evidential arguments have their effect on the explanations.

A new, admittedly prejudiced, generalisation can be added to T_r in order to make explicit one of the reasons that the women's story is more plausible.

ge₄: *x* has previous convictions $\Rightarrow_E$ *x* has gun

Say that there is evidence for the fact that Rijkbloem has previous convictions. Now the explanation S_r is even better because now Rijkbloem has gun is supported by an evidential argument; the abductively inferred part of S_r, the "gap" in the story so to speak, was Rijkbloem has a gun. This proposition can now be inferred from evidence using **ge₄**. However, it can be argued that generalisation **ge₄** is based on based on prejudices and an argument can be constructed for the conclusion that the generalisation is not valid.

Another intuitive reason for the fact that the womens' story is more plausible concerns the generalisation **gc₆**. It can be argued that this generalization is not valid, on the grounds that it is highly improbable

For the sake of the example, we return to the situation from the beginning of this section: the two explanations S_r and S_m from Figure 4 without the undercutting arguments for the fact that Rijkbloem wants to protect himself. We present some new observations:

O_n = {Rijkbloem's gun was of type A,
mother's gun was of type B,
¬(bullet casings on crime scene),
¬(gunpowder on Rijkbloem's hands),
Rijkbloem was 50 centimeters away from father,
mother was 2 meters away from father,
the angle at which the bullet entered mr. Lammers is θ_2}

These observations all follow from evidence: the women state that Rijkbloem shot with a "black pistol, similar to the one the police carries". Rijkbloem testifies that the mother's gun was "a small, revolver-like pistol". The police report states that there were no bullet casings found on the crime scene and that here was no powder on Rijkbloem's hands. A forensics report states that the angle at which the bullet entered was θ_1. Assume that all the conclusions in O_n follow from defensible arguments and $O_u = O \cup O_n$. The explananda F_u are now changed into {father dies, the angle at which the bullet entered mr. Lammers is θ_2¬(bullet casings on crime scene), ¬(gunpowder on Rijkbloem's hands)}. The causal theory T is also updated with some new generalisations:

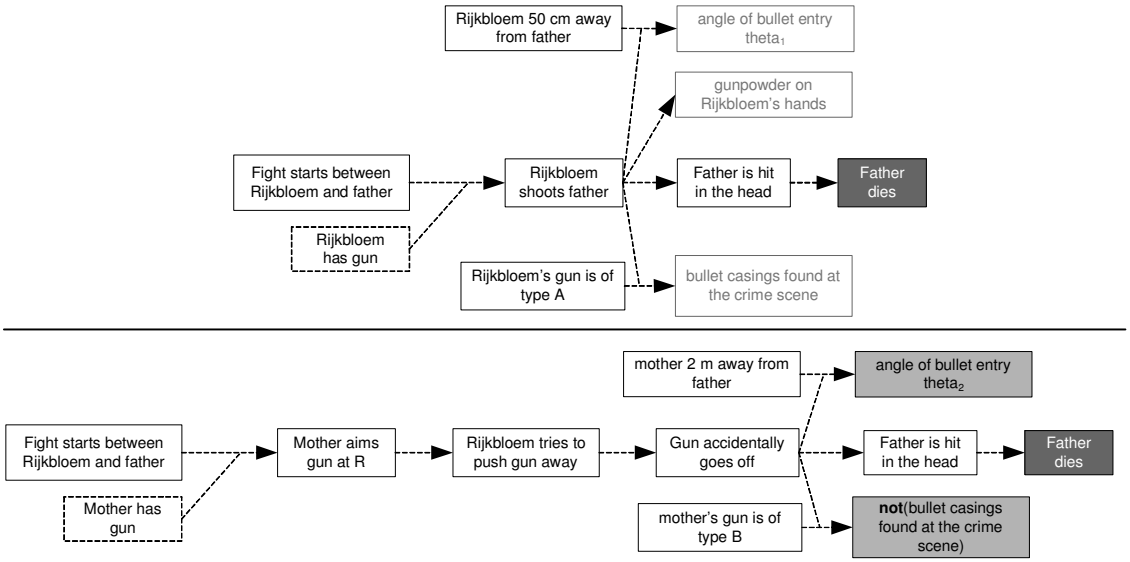

Figure 5: the updated explanations S_{ru} and S_{mu}

gc$_9$: *x* shoots *y* ∧ *x* is 50 centimeters away from *y*
⇒$_C$ the angle at which the bullet entered *y* is θ_1
gc$_{10}$: *x* shoots *y* ∧ *x* shoots with gun type A
⇒$_C$ bullet casings on crime scene
gc$_{11}$: *x* shoots *y* ⇒$_C$ *x* has gunpowder on his hands
gc$_{12}$: *x*'s gun accidentally goes off ∧ *x* is 2 meters away from *y*
⇒$_C$ the angle at which the bullet entered *y* is θ_2
gc$_{13}$: *x*'s gun accidentally goes off ∧ *y*'s gun is of type type B
⇒$_C$ ¬(bullet casings on crime scene)

Note that these generalisations are typically generalisations that follow from expert statements; for example, an expert on guns testifies that the semi-automatic pistol with which Rijkbloem shot ejects its empty casings when fired. Arguments about the validity and quality of generalisations can play a big part in determining the right causal theory. However, in this paper, we do not discuss this in great detail.

Two new explanations can now be built (Figure 5). The first is an updated version of the women's story: $S_{ru} = T_{ru} \cup H_{ru}$, where T_{ru} = {**gc$_1$**, **gc$_2$ gc$_3$**, **gc$_9$**, **gc$_{10}$ gc$_{11}$**} and H_r = {fight started between Rijkbloem and father, Rijkbloem had gun, Rijkbloem was 50 cm away from father, Rijkbloem's gun was of type A}. This explanation explains the explanandum father dies. However, it also explains three propositions which are contradicted by defensible arguments; the angle at which the bullet entered father is θ_2 instead of θ_1, Rijkbloem did not have gunpowder on his hands and there were no bullet casings found at the crime scene. The propositions that contradict an argument's conclusion have been rendered in a lighter colour in Figure 5.

The second new explanation is an updated version of Rijkbloem's story: $S_{mu} = T_{mu} \cup H_{mu}$ where T_{mu} = {**gc$_4$**, **gc$_5$**, **gc$_6$**, **gc$_7$**, **gc$_{12}$**, **gc$_{13}$**} and H_m = {fight started between Rijkbloem and father, mother had gun, mother was 2 m away from father, mother's gun was of type B}. This explanation does not contradict any conclusions of arguments. It does explain two explananda which follow from defensible arguments, about the angle of entry and about the fact that no bullet casings were found (these propositions have been rendered as dark grey boxes in Figure 5).

So it seems that, while Rijkbloem's story fares less well on plausibility, it fares better in light of the evidence. This is not to say Rijkbloem did not shoot the father. However, the analysis does show that "good" stories are not necessarily true, and "bad" stories are not necessarily untrue

6. CONCLUSION

In this paper we have argued for a combined story and argumentative approach to reasoning with evidence. In our opinion, such an approach solves the problems of the individual approaches as discussed in section 3.2. When stories are represented as causal networks, the overview of "what happened" in the case is retained, and the abductive approach to explaining the explananda nicely captures the intuitive process of comparing different stories. Also, the role of stories as "gap-fillers" is captured by this abductive reasoning.

Adding evidential arguments to connect the sources of evidence solves the problems the abductive approach has with these sources, and gives the individual pieces of evidence a separate place in the theory. The conclusions of the evidential arguments can be used to define criteria for comparing explanations, and the evidential arguments are a good tool for analysing the causal generalisations in a theory.

As for future research, we did not discuss alternative story structures like intentional structures (such as the one used by Pennington and Hastie [12]) and content-based structures (such as the story schemes used by Schank and Abelson in their seminal work on story understanding [17]). However, these other story structures also have a causal model as their basis; the schemes are added as an extra layer to this causal structure. So our basic causal approach allows us to extend our theory, adding other layers as necessary.

Another limitation of the present work is that we take a static viewpoint: we provide an analysis framework for the current status of an argumentative and story-based analysis of the evidence about case facts, and postpone attention to the dynamics of developing and refining such an analysis. We trust that our work can be a fruitful basis for further research in which

the dynamics of proof analysis is given due attention. In our opinion, our approach can be used to explicate heuristics for police investigation into the case facts, using alternative explanations to steer the further gathering of evidence.

ACKNOWLEDGMENTS

This research was supported by the Netherlands Organisation for Scientific Research (NWO) under project number and 634.000.429. Henry Prakken was also partially supported by the EU under IST-FP6-002307 (ASPIC).

REFERENCES

[1] Anderson, T. J., Schum, D. A., and Twining, W. L. (2005) *Analysis of Evidence*, 2nd edition, Cambridge University Press, Cambridge.

[2] Bex, F. J., Prakken, H., Reed, C., and Walton, D. N. (2003) Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law* 11, pp. 125 - 165.

[3] Bex, F.J., Prakken, H. and Verheij, B. (2006) Anchored Narratives in Reasoning about Evidence. In T.M. van Engers (ed.), *Legal Knowledge and Information Systems. JURIX 2006: The Nineteenth Annual Conference*, pp. 11-20, IOS Press, Amsterdam.

[4] Van den Braak, S.W., & Vreeswijk, G.A.W. (2006), AVER: Argument visualization for evidential reasoning. In T.M. van Engers (ed.), *Legal Knowledge and Information Systems. JURIX 2006: The Nineteenth Annual Conference*, pp. 151, 156, IOS Press, Amsterdam.

[5] Crombag, H. F. M., Wagenaar, W. A., and Van Koppen, P. J. (1994) *Dubieuze Zaken: de psychologie van strafrechtelijk bewijs*, 2nd edition, Contact, Amsterdam. English edition: Wagenaar, W.A., Koppen, P.J. van & Crombag, H.F.M. (1993), *Anchored Narratives. The Psychology of Criminal Evidence*. London: Harvester Wheatsheaf.

[6] Dung, P.M. (1995), On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. *Artificial Intelligence* 77, pp. 321-357.

[7] Hage, J. C. (1997). Reasoning with Rules. *An Essay on Legal Reasoning and Its Underlying Logic*. Kluwer Academic Publishers, Dordrecht.

[8] Hage, J. and Verheij, B. (1999). The law as a dynamic interconnected system of states of affairs: a legal top ontology. *International Journal of Human-Computer Studies* 51, pp.1043 - 1077.

[9] Keppens, J. and Schafer, B. (2006) Knowledge based crime scenario modelling. *Expert Systems with Applications* 30: pp. 203 - 222.

[10] Lucas, P. (1997). Symbolic diagnosis and its formalisation. *The Knowledge Engineering Review* 12, pp. 109 - 146.

[11] Pearl, J. (1988) Embracing Causality in Default Reasoning. *Artificial Intelligence* 35, pp. 259 - 271.

[12] Pennington, N. and Hastie, R. (1993) The story model for juror decision making. In Hastie, R. (eds.), *Inside the Juror, The Psychology of Juror Decision Making*, Cambridge University Press, Cambridge.

[13] Pollock, J. (1995) *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press, Cambridge, MA.

[14] Poole, D. (1998) Learning, Bayesian Probability, Graphical Models, and Abduction, In Flach, P. and Kakas, A.(eds), *Abduction and Induction: essays on their relation and integration*, Kluwer, 1998.

[15] de Poot, C. J., Bokhorst, R. J., Van Koppen, P. J., and Muller, E. R. (2004) *Rechercheportret - Over dilemma's in de opsporing* (*Investigation portrait – about dilemmas in investigation*), Kluwer, Alphen a.d. Rijn.

[16] Prakken, H. and Vreeswijk, G. (2002) Logical systems for defeasible argumentation. In D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, second edition, Vol 4, pp. 219-318. Kluwer Academic Publishers, Dordrecht etc.

[17] Schank, R. C., and Abelson, R. P. (1977) *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*, Lawrence Erlbaum, Hillsdale, NJ.

[18] Thagard, P. (2005) Testimony, Credibility and Explanatory Coherence. *Erkenntnis* 63, pp. 295 – 316.

[19] Thagard, P. and Shelley, C. P. (1997). Abductive reasoning: Logic, visual thinking, and coherence. In M.-L. Dalla Chiara et al. (Eds.), *Logic and scientific methods*. pp. 413-427, Kluwer, Dordrecht.

[20] Tillers, P. and Schum, D. A. (1991) A Theory of Preliminary Fact Investigation. *University of California Davis Law Review* 24, pp. 931 - 1012.

[21] Toulmin, S. E. (2003) *The Uses of Argument*, Updated Edition, in 1958), Cambridge University Press, Cambridge.

[22] Twining, W. L. (1994) *Rethinking Evidence: Exploratory Essays*, Northwestern University Press, Evanston, IL

[23] Verheij, B. (2003) Dialectical argumentation with argumentation schemes: An approach to legal logic. *Artificial Intelligence and Law* 11, pp.167 - 195.

[24] Verheij, B. (2005) *Virtual Arguments. On the Design of Argument Assistants for Lawyers and Other Arguers*. T.M.C. Asser Press, The Hague.

[25] Walton, D.N. (1996) *Argumentation Schemes for Presumptive Reasoning*, Lawrence Erlbaum Associates, Mahwah, N.J.

[26] Wigmore, J. H. (1931) *The Principles of Judicial Proof or the Process of Proof as Given by Logic, Psychology, and General Experience, and Illustrated in Judicial Trials*, 2nd edition, Little, Brown and Company, Boston, MA