

# An integrated theory of causal stories and evidential arguments

Floris Bex

Department of Information and Computing Sciences  
Utrecht University, The Netherlands  
f.j.bex@uu.nl

## ABSTRACT

In the process of proof alternative stories that explain ‘what happened’ in a case are tested using arguments based on evidence. Building on the author’s earlier hybrid theory, this paper presents a formal theory that combines causal stories and evidential arguments, further integrating the different types of reasoning in a framework for structured argumentation. This then allows for correct reasoning with causal and evidential rules, and further integrates arguments and stories by grounding them both in well-known dialectical argumentation semantics.

## 1. INTRODUCTION

The process of legal proof is one of inference to the best explanation (IBE), in which alternative stories or scenarios that explain ‘what happened’ in a case are tested and compared using arguments based on evidence. This combination of explanatory, story-driven and argumentative, evidence-driven reasoning is captured in the *hybrid theory* [4], which combines theories of causal-abductive diagnosis [15, 10] with a framework for structured argumentation [3].

In the hybrid theory, causally coherent stories that explain the evidence can be abductively inferred: if we observe some effect  $e$  (e.g. a body with gunshot wounds) we are allowed to hypothesise a possible cause  $c$  (e.g. the person was shot). Stories can then be supported or attacked by arguments based on evidence. For example, a witness who saw that the victim was shot provides evidence to support our hypothesis that the victim died because he was shot, while a coroner’s report stating that the victim was already dead when he sustained the gunshot wounds attacks the hypothesis. Such counterarguments can point to alternative explanations: maybe someone wanted it to look like the victim was shot to cover up some other cause of death. Ultimately, the alternative stories have to be compared according to how much of the evidence supports and attacks them.

The hybrid theory is significantly more expressive than most logical frameworks for IBE [15, 10], which were orig-

inally intended for automated diagnosis within relatively small and pre-defined domains. It allows for reasoning with both causal rules ( $c$  causes  $e$ ) and evidential rules ( $e$  is evidence for  $c$ ) and reasoning about the elements of stories, such as the events and the causal links that comprise them. Furthermore, where classical theories of IBE compare stories on basic criteria such as the minimum number of assumptions, the hybrid theory also includes other more fine-grained story comparison criteria based on the number of supporting or attacking arguments.

The hybrid theory stops short of truly *integrating* causal, story-based reasoning and evidential argumentation. While there is a clear link between arguments and stories — arguments can have elements of stories as their conclusion — arguments and stories remain completely separate entities, each with their own definitions related to conflict and comparison. Thus, two contradictory arguments attack each other but two contradictory stories do not, and two alternative stories are in competition, whilst two arguments that present alternative explanations are not. Furthermore, the status of arguments is determined using the standard semantics of Dung [8], whereas the status of stories is determined with more absolute measures (i.e., by counting the number of supporting and attacking arguments based on evidence [4]). The acceptability of a certain claim thus depends on whether it is modelled as part of a story or an argument.

The outlines of the integrated theory were discussed in a previous position paper [2], in which both causal-abductive and evidential argumentation were captured in a structured argumentation framework similar to  $ASPIC^+$  [11]. In this paper, this integrated framework is further extended and formalised as an instance of  $ASPIC^+$ . This new integrated theory allows for causal and evidential arguments as well as mixed causal-evidential arguments, while still adhering to the constraints imposed by Pearl’s C-E system [12]. Furthermore, the integrated theory allows for a natural rendering of the different modes of reasoning employed in the process of proof, as will be demonstrated with a small case-study that is used as a running example.

The case that is used concerns that of Danny Rijkbloem [23, 4], who has a considerable list of sentences (theft, robbery) starting when he was 15 years old. Nicole Lammers is a baker’s daughter who had a relationship with Rijkbloem and lived together with him. At some point Nicole decided under pressure of her parents to break up with Rijkbloem and leaves him. A few days after the breakup, Nicole and her parents went to Rijkbloem’s house to pick up some of Nicole’s stuff and got into an argument with Rijkbloem. At

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

ICAIL ’15, June 08 - 12, 2015, San Diego, CA, USA

Copyright is held by the author. Publication rights licensed to ACM.

ACM 978-1-4503-3522-5/15/06...\$15.00.

DOI: <http://dx.doi.org/10.1145/2746090.2746094>.

some point, a scuffle developed, which ended in the father getting shot in the head. When the police, who had been informed by Rijkbloem, arrived, the father was already dead. 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. 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.

## 2. A HYBRID THEORY OF STORIES AND ARGUMENTS

Two influential methods for the analysis of evidence are the argument-based approach [3], which focuses on arguments based on evidence, and the story-based approach [23, 14], which uses hypothetical stories or scenarios to explain the evidence. In the argument-based approach the reasoning goes from evidence to conclusions by the application of evidential rules (*e is evidence for c*), and arguments for and against the conclusion are considered and compared. In the story-based approach multiple scenarios consisting of causal rules (*c is a cause for e*) are constructed to explain the evidence, and the scenarios are considered and compared.

Both the evidential-argument approach and the causal-story approach have advantages. An inference from a piece of evidence (e.g. a witness testimony) to a conclusion is best captured using evidential rules [4, 7], and arguments based on evidence can be compared using Dung’s solid and intuitive framework [8]. However, certain parts of a case (such as the cause of death, or the behaviour of the suspect) are more easily captured using causal rules [7]. Furthermore, scenarios help in providing an overview of the case as they are understandable stories that explain the evidence in a natural way [14]. What is missing from purely causal, story-based approaches, however, is the possibility to talk about elements (events, causal links) in a scenario using evidence.

The combination of evidential arguments and causal scenarios then seems to be an intuitive and analytically useful perspective for looking at cases and evidence. Hence, we proposed the hybrid theory [4], a combination of formal, causal story-based reasoning (see e.g. [10, 15]) and formal

evidential argumentation [3].

The hybrid theory consists of a set of evidence, a set of hypotheses and a set of defeasible inference rules, both causal (*c causes e*) and evidential (*e is evidence for c*). Scenarios are built by taking some observations and abductively inferring hypotheses. So given an observation *o* and a causal rule '*c causes o*', we can abductively infer *c*. Arguments can be built by taking evidence and consecutively applying evidential inference rules to infer conclusions. So given a piece of evidence *e* and an evidential rule '*e is evidence for c*' we can conclude *c*. The connection between arguments and stories lies in the conclusions of arguments. These conclusions can be elements of a story (i.e., the evidence on which the argument is based *supports* the story) or the negation of elements of a story (i.e., the evidence on which the argument is based *contradicts* the story).

Figure 1 shows the Rijkbloem case as a combination of arguments and stories, where arrows with open arrowheads stand for evidential rule applications, and arrows with closed arrowheads stand for causal rule applications. The main observation that needs to be explained in this case is **Father died**. Now, if we assume the hypothesis **Rijkbloem and father argued**, we can build two stories *S<sub>r</sub>* and *S<sub>m</sub>* that explain this observation. Story *S<sub>r</sub>* says that Rijkbloem shot father is the cause of **Father was hit**, while story *S<sub>m</sub>* argues that **Father was hit** was caused by **Mother shot father**, which itself was caused by **Rijkbloem pushed mother**.

Stories *S<sub>r</sub>* and *S<sub>m</sub>* are *alternative explanations* for the same facts. What this means is that, in absence of any explicit links between them, we consider the chances of both explanations being true at the same time being very low (i.e., either Rijkbloem shot father or mother accidentally shot father, but not both). In Pearl’s terms [12], two causes of a common consequence *compete* with each other. Thus, accepting one explanation will severely diminish our belief in the other. In the literature on Bayesian networks [13, 20], this phenomenon is called ‘explaining away’. In logical accounts of abductive reasoning like the hybrid theory, this can be enforced by only accepting the simplest explanations, that is, those that are (subset) minimal in terms of the hypotheses and rules they use [15, 16, 10].

The evidence is rendered as grey boxes in figure 1. From the evidence, elements of the stories can be inferred using evidential reasoning. For example, from **Women testimony<sub>3</sub>**, a statement by Nicole and mother that it was Rijkbloem

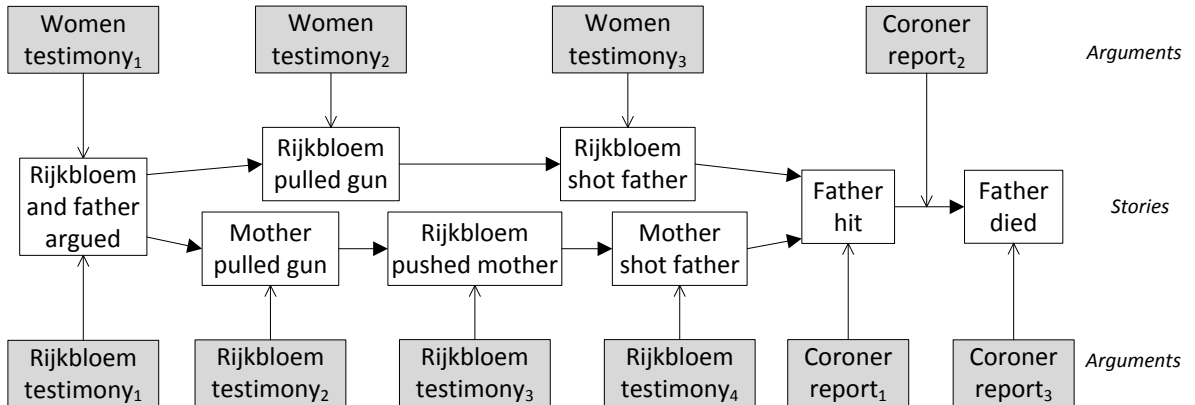


Figure 1: The two Rijkbloem scenarios as a hybrid case

who shot father, we can infer the conclusion that **Rijkbloem shot father** because the testimony is evidence for the fact that the event happened. It is also possible to support or attack (applications of) causal inference rules in the stories. For example, the evidence **Coroner Report<sub>2</sub>**, an autopsy report by the coroner that father was killed by a gunshot to the head, allows us to conclude that there is a causal link between **Father hit** and **Father died**.

Given alternative scenarios such as  $S_r$  and  $S_m$ , the question is how to compare them. In the hybrid theory, an important role is played by *evidential support*, the evidence that supports a scenario because it is part of the scenario or because it is a premise of an argument that has as its conclusion an element in the scenario. Similarly, another important concept is *evidential contradiction*, evidence that contradicts a scenario because it is a premise of an argument that contradicts the scenario. These two concepts can be used to compare scenarios; for example, we can look how much of the total evidence in a case supports or attacks a story, or we can see whether a story is supported or attacked by an important piece of evidence, thus getting an indication of the relevance and coverage of a scenario [14]. Note that the status of arguments, that is, whether they are part of an extension [8] may influence these comparison criteria [4]. For example, we could say that only evidence which is part of an argument in an extension counts towards the evidential support or contradiction.

## 2.1 Stories and arguments: two sides of the same coin

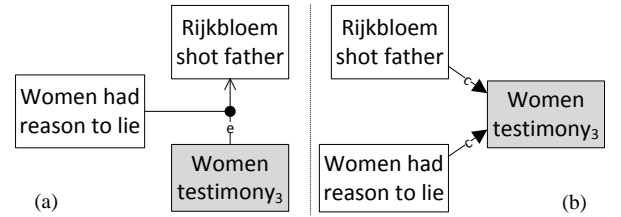
The two separate approaches to reasoning with evidence, arguments and stories, often blend into each other, both in the literature and in real cases. For example, the Anchored Narratives Theory [23] has clear argumentative aspects, as it is possible to ‘argue’ for a particular position in a case using causal stories instead of evidential arguments. This is because ultimately, it is not the structure of the reasoning (i.e. evidential, causal) but rather the intention of the speaker that determines whether a story is used to, for example, explain or argue [5]. This is already evident in the wording chosen earlier in this paper: story  $S_2$  *argues* that father was shot accidentally.

Further overlap between the two parts of the hybrid theory concerns the fact that causal and evidential reasoning are closely entwined: if we have a causal rule  $c$  *causes*  $e$  then we will usually also accept that, in some way,  $e$  *is evidence for*  $c$  [12]. For example, father getting shot can cause him to die, so the observation that father died can be seen as evidence for the fact that he was shot. This can also be true the other way: if we have an evidential rule, we will often also accept some sort of causal relation, precisely because the evidential rule is based on a causal rule. However, this depends on our interpretation of the evidential rule  $e$  *is evidence for*  $c$ . If we interpret the rule broadly, as  $e$  being any sign that increases our belief in  $c$ , it need not be based on a causal relation. For example [24]: dark clouds are evidence for the fact that it will rain later, but we would not say that the fact that it will rain later causes dark clouds. However, if we interpret the meaning of  $e$  *is evidence for*  $c$  as  $e$  being evidence for some event  $c$  that happened earlier, then we are quite safe in assuming that  $e$  was caused (at least partly) by  $c$ .

The close relation between causal and evidential reasoning is also important for formal models: Poole [15, 16] argues

that information in causal models can equally well be represented by evidential rules; instead of abductive reasoning we then perform modus-ponens style reasoning (‘deduction’) to infer the hypotheses based on our observations. So, for example, instead of abductively inferring **Rijkbloem shot father** from the observation **Father was hit** and a causal rule ‘**Rijkbloem shot father** *causes* **Father was hit**’, we use the observation **Father was hit** and the evidential rule ‘**Father was hit** *is evidence for* **Rijkbloem shot father**’ to infer the conclusion **Rijkbloem shot father**.

The relation between causal and evidential reasoning is also evident in the different ways in which, for example, the inferences and attacks based on witness testimony can be captured. One way to do this is to use an evidential rule, like in the hybrid case in figure 1. An argument based on this rule can be undercut (attacked) by claiming that, for example, the women had a good reason to give false statements. This way of modelling inference based on witness testimony is captured in figure 2a, where the arrow with the open arrowhead is again an evidential inference and the arrow with the circle at the end is an attack relation. A second way to model testimonies is by regarding a testimony as being *caused* either by the event to which it testifies, or by other causes, such as a reason to lie. There are then two alternative abductive explanations for **Women testimony<sub>3</sub>** (figure 2): **Rijkbloem shot father** and **Women had reason to lie**, and accepting one will deny the other. This link between explaining away and undercutting has recently been formalised in [20], who show that a Bayesian network with two competing causal explanations can be translated into two attacking arguments.



**Figure 2: Witness testimony as arguments (a) or stories (b)**

The above examples show that causal stories and evidential arguments are essentially two sides of the same coin: in theory any case can be modelled using either causal stories or evidential arguments. With respect to causal and evidential, it is in some ways arbitrary how we should capture a relation. Pearl [12] argued that people have a hard time expressing their knowledge just as either causal or evidential rules, and van den Braak et al. [7] found that, while there are situations in which people consistently choose either causal or evidential modelling techniques, there are also many examples where people interpret a relation evidentially or causally in equal measure.

Similarly, some aspects of a case will be more conveniently modelled in terms of arguments and others in terms of stories. For example, if the decision whether the suspect is guilty or not hinges on a single fact (e.g. whether the suspect was at a particular location at a particular time), alternative stories are less important and the main analytical technique is to carefully consider all the arguments for and against this fact. In other cases, such as cases where there is no di-

rect evidence for the main event, it is important to consider the alleged events as a coherent story in order to be able to decide on the basis of more circumstantial evidence.

In sum, it is difficult to say that there is one ‘correct’ way of reasoning in the process of proof. The exact choice whether to use causal rules, evidential rules, stories or arguments is often determined by personal preference as much as features of the case. Hence, we need a theory that captures and integrates both types of reasoning. While the hybrid theory is a good first step, there are some problems, as explained in the next section.

## 2.2 Problems for the hybrid theory

In an integrated theory of causal stories and evidential arguments, it should be possible to combine reasoning with both evidential and causal rules in a correct way. For example, in the Rijkbloem case the two women testified that Rijkbloem used a ‘black pistol’. From this we can evidentially infer that Rijkbloem probably shot father with a type of pistol that usually ejects its cartridges when fired. We can then predict a so-called ‘story consequence’: if the story about the type of pistol is really true, then we should find empty cartridges at the scene of the crime. Figure 3 shows this line of reasoning: from **Women testimony<sub>3</sub>** we evidentially infer **Rijkbloem shot father** using an evidential rule, and we then subsequently causally infer **Casings at crime scene** from **Rijkbloem shot father** using a causal rule. While these types of mixed causal-evidential (C-E) arguments are very common in the process of proof, they cannot be constructed in the hybrid theory.

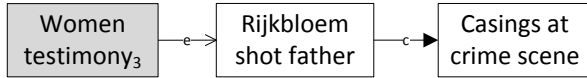


Figure 3: Mixed evidential and causal reasoning

Whilst mixed C-E reasoning is thus a desirable feature of a theory for reasoning in the process of proof, care must be taken that only correct or valid lines of reasoning are allowed [12]. Take, for example, the reasoning in figure 4. Here, we first causally infer **Father hit** from **Rijkbloem shot father**, after which we infer the explanation **Mother shot father** from **Father hit** using an evidential reasoning. By itself, the evidential step from **Father hit** to **Mother shot father** is fine, as the latter is a possible cause of the former. However, the problem here is that we have already assumed that **Rijkbloem shot father**, and we should not be able to infer the alternative, competing cause for **father hit** (i.e. **Mother shot father**) from this. In the hybrid theory, this is not a problem: mixed arguments are not possible, and if there are two different causes they will be part of alternative explanations due to the subset minimality requirement of explanations. However, an integrated theory that does allow for mixed C-E reasoning will have to explicitly take this into consideration.

An integrated theory should also allow us to emphasize argument-oriented elements or story-oriented elements in a case analysis without changing the essential conclusions. However, in the current hybrid theory this is not the case. Take, for example, the two alternative causes for **F was hit in the head** as two (simple) causal stories  $S'_r$  and  $S'_m$  (figure 5a). If we change the causal rules used into evidential

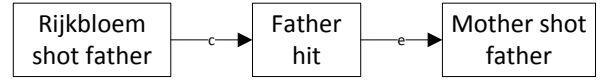


Figure 4: Invalid mixed evidential and causal reasoning

rules, we get two arguments  $A_r$  and  $A_m$  that do not attack each other (figure 5b), even though the stories  $S'_r$  and  $S'_m$  are in competition.

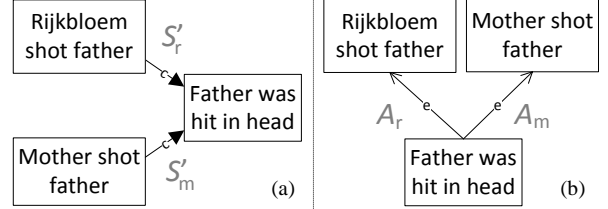


Figure 5: Alternative stories (a) or arguments (b)

Another problem is that the comparison of alternative scenarios and the comparison of attacking arguments is defined differently, and that thus the acceptability of a claim may thus depend on whether it is modelled as part of a story or an argument. For example, an argument  $B$  with the conclusion  $\neg(\text{Rijkbloem shot father})$  attacks story  $S'_r$  (as well as the main story  $S_r$ ), adding to its evidential contradiction.  $B$  will also attack argument  $A_r$ , but since determining the ‘winning’ argument is done differently than determining the ‘winning’ story, it is not guaranteed that a case with  $S'_r$  and  $B$  has the same outcome as a case with  $A_r$  and  $B$ .

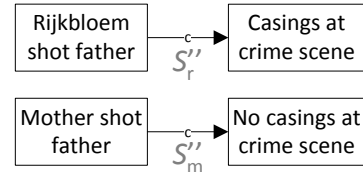


Figure 6: Two contradictory stories

Related to this is the problem that contradictory stories do not attack each other. For example, Rijkbloem argued that mother used a ‘small revolver’ of the type that does *not* eject empty cartridges when shot. Thus, we can build two small causal stories that explain contradictory observations (figure 6):  $S''_r$  predicts that there should be shell casings at the crime scene whilst  $S''_m$  predicts that there should not be shell casings at the crime scene. According to the definitions of the hybrid theory, these two stories do not contradict each other, even though they are clearly incompatible. Thus, it is not possible to capture, for example, the idea of an alibi story in the hybrid theory, which is not an alternative explanation (it does not explain the main questions in a case, e.g. why the victim died) but it does contradict the prosecution’s story (that the suspect killed the victim).

## 3. AN INTEGRATED THEORY

The above discussion indicates that we need to reconsider the relation between arguments and scenarios: we need to

be able to reason with alternative explanations even if they are modelled as evidential arguments and we need to be able to consider contradictory causal scenarios. Furthermore, alternative and contradictory claims should be defined as incompatible irrespective of whether they are part of an argument or a story, and their evaluation should always follow the same rules. Finally, both evidential and causal inference should be allowed in an argument, so that we have a fully integrated theory instead of a hybrid of causal, scenario-based and evidential argumentative reasoning.

In this section, the integrated theory for causal stories and evidential arguments is presented. This integrated theory is explicitly based on the *ASPIC*<sup>+</sup> framework [11]. As discussed in section 2.1 (see also [2]), the formal structure of arguments and stories is very similar: both are derivations in some defeasible logic, the difference being that scenarios use causal rules and are based on hypotheses and arguments use evidential rules and are based on evidence. This means that we need at least a logical language, causal and evidential inference rules and a knowledge base of hypotheses and evidence, all of which are captured by the *ASPIC*<sup>+</sup> notion of an argumentation system.

**DEFINITION 1. [Argumentation systems]** An *argumentation system* is a triple  $AS = (\mathcal{L}, \mathcal{R}, n)$  where:

- $\mathcal{L}$  is a logical language closed under negation ( $\neg$ ).
- $\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_d$  is a set of strict ( $\mathcal{R}_s$ ) and defeasible ( $\mathcal{R}_d$ ) inference rules, where  $\mathcal{R}_d = \mathcal{R}_c \cup \mathcal{R}_e$  is a set of causal ( $\mathcal{R}_c$ ) and evidential ( $\mathcal{R}_e$ ) defeasible inference rules and  $\mathcal{R}_s \cap \mathcal{R}_d = \emptyset$ .
- $n : \mathcal{R}_d \rightarrow \mathcal{L}$  is a naming convention for defeasible rules.

We write  $\psi = -\varphi$  just in case  $\psi = \neg\varphi$  or  $\varphi = \neg\psi$ .

### 3.1 Causal and evidential inference rules

Definition 1 is similar to the original *ASPIC*<sup>+</sup> definition save for the addition of sets  $\mathcal{R}_c$  and  $\mathcal{R}_e$  for causal and evidential rules. The rules can be defined as follows.

**DEFINITION 2. [Inference Rules]** Let  $\varphi_i, \psi$  be elements of  $\mathcal{L}$ .

- A *strict rule* is of the form  $\varphi_1, \dots, \varphi_n \rightarrow \psi$ . Informally, if  $\varphi_1, \dots, \varphi_n$  holds then without exception  $\psi$  holds.
- A *causal defeasible rule* is of the form  $\varphi \Rightarrow_c \psi$ . Informally, if  $\varphi$  holds then this has (presumably) caused  $\psi$ .
- An *evidential defeasible rule* is of the form  $\varphi \Rightarrow_e \psi$ . Informally, if  $\varphi$  holds then (presumably) this is evidence for  $\psi$ .

For the strict rules it is assumed that the common inference rules of deductive logic are in  $\mathcal{R}_s$ . Note that whereas strict rules can have more than premise, causal and evidential rules have only one wff from  $\mathcal{L}$  as a premise. This is a change from *ASPIC*<sup>+</sup> which was made to simplify several definitions in the integrated theory, most notably definition 3, which compares premises and conclusions of causal and evidential rules. This does not diminish the expressivity of the integrated framework: multiple premises can be combined into one conjunction using a strict inference rule, and this conjunction can then serve as the single premise of a causal or evidential rule.

In the Rijkbloem case, we have the following causal inference rules (where *R* is Rijkbloem, *F* is father, *M* is mother and *g* is gun).

- $rc_1$  *R and F argued*  $\Rightarrow_c$  *R pulled g*
- $rc_2$  *R pulled g*  $\Rightarrow$  *R shot F*
- $rc_3$  *R shot F*  $\Rightarrow_c$  *F hit*
- $rc_4$  *R and F argued*  $\Rightarrow_c$  *M pulled g*
- $rc_5$  *M pulled g*  $\Rightarrow_c$  *R pushed M*
- $rc_6$  *R pushed M*  $\Rightarrow_c$  *M shot F*
- $rc_7$  *M shot F*  $\Rightarrow_c$  *F hit*
- $rc_8$  *F hit*  $\Rightarrow_c$  *F dies*
- $rc_9$  *R shot F*  $\Rightarrow_c$  *casings*
- $rc_{10}$  *M shot F*  $\Rightarrow_c$   $\neg(\text{casings})$

There are also multiple evidential rules in the case, some examples of which are as follows.

- $re_1$  *Rijkbloem testimony*<sub>4</sub>  $\Rightarrow_e$  *M shot F*
- $re_2$  *Women testimony*<sub>3</sub>  $\Rightarrow_e$  *R shot F*
- $re_3$  *Coroner report*<sub>1</sub>  $\Rightarrow_e$  *F hit*
- $re_4$  *Coroner report*<sub>2</sub>  $\Rightarrow_e$   $rc_3$
- $re_5$  *Women lied*  $\Rightarrow_e$   $\neg re_2$
- $re_6$  *F hit*  $\Rightarrow_e$  *R shot F*
- $re_7$  *F hit*  $\Rightarrow_e$  *M shot F*

Note that causal as well as evidential rules can be supported - rule  $re_4$ , for example, allows us to infer  $rc_3$  from *Coroner report*<sub>2</sub> - as well as attacked - rule  $re_5$ , for example, denies rule  $re_2$  (see figure 2.<sup>1</sup>).

It is important to note that there is an implicit causal theory underlying the causal and evidential rules, that is, the domain can sensibly be thought of in terms of causes and effects (cf. [16]). Hence, we assume that a causal rule *c causes e* is based on a causal relation this underlying theory and that if we say that *e is evidence for c* there is a causal relation between *c* and *e* in the underlying theory. The causal theory has only one constraint, namely that it does not allow for causal cycles: if *a* is a cause of *b*, then *b* cannot also be a cause of *a*. Causality is a contentious topic and it is very easy to disagree about whether one event causes another, or whether one event is evidence for another. The causal theory does not enforce exactly which causal and evidential rules are ‘correct’ and which are not - disagreement about rules can be expressed by counterarguments to causal rules.

### 3.2 Capturing alternative causes

When modelling the causal and evidential relations based on an underlying causal theory, it is important to consider the definition of *alternative causes*. Like in the hybrid theory, two different causes with the same consequences are considered to be alternative explanations. As was discussed in section 2.2, it is also possible to have two alternative causes in evidential arguments, because an evidential rule is essentially nothing more than an explicit expression of an abductive reasoning step applied to a causal rule. Thus, evidential alternatives are exactly the opposite of causal alternatives: two different conclusions that are evidentially inferred from the same premises are considered as alternatives (see figure 5). Finally, it is also possible to have alternative causal-evidential arguments: if some conclusion *e* is causally

<sup>1</sup>For simplicity, no distinction is made between a rule being inapplicable in all cases and a rule being inapplicable in only some cases

inferred from a cause  $c_1$  and another cause  $c_2$  is then evidentially concluded from  $e$ , causes  $c_1$  and  $c_2$  are alternatives.

**DEFINITION 3. [Alternatives]** Causes  $\varphi$  and  $\chi$  are *alternatives* if for any pair of rules  $r_i, r_j \in \mathcal{R}_c \cup \mathcal{R}_e$ .

1.  $r_i = \varphi \Rightarrow_c \psi$  and  $r_j = \chi \Rightarrow_c \psi$  and  $\varphi \neq \chi$ ; or
2.  $r_i = \psi \Rightarrow_e \varphi$  and  $r_j = \psi \Rightarrow_e \chi$  and  $\varphi \neq \chi$ ; or
3.  $r_i = \varphi \Rightarrow_c \psi$  and  $r_j = \psi \Rightarrow_e \chi$  and  $\varphi \neq \chi$ .

We can see that **R shot F** and **M shot F** are alternatives because for any pair of the following rules either definition 3 (1), (2) or (3) holds.

```
rc3  R shot F  $\Rightarrow_c$  F hit
rc7  M shot F  $\Rightarrow_c$  F hit
re6  F hit  $\Rightarrow_e$  R shot F
re7  F hit  $\Rightarrow_e$  M shot F
```

Note that in the hybrid theory, alternatives are not explicitly defined. Rather, the idea of definition 3(1) is implicitly enforced by comparing only subset-minimal causal explanations of a consequence. Furthermore, evidential or C-E alternatives as defined in (2) and (3) are not distinguished in the hybrid theory.

### 3.3 Constructing arguments

Arguments and stories in the integrated theory are based on a knowledge base. Like in the hybrid theory, this knowledge base contains both the evidence and the possible hypotheses.

**DEFINITION 4. [Knowledge bases]** A *knowledge base* in an  $AS = (\mathcal{L}, \mathcal{R}, n)$  is a set  $\mathcal{K} \subseteq \mathcal{L}$  consisting of two disjoint subsets  $\mathcal{K}_e$  (the *evidence*) and  $\mathcal{K}_h$  (the *hypotheses*).

The evidence in  $\mathcal{K}_e$  is similar to the axiom premises in  $AS-PIC^+$  in that they cannot be sensibly denied or attacked. What this means is not that, for example, the veracity of a witness testimony cannot be questioned, but rather that the *existence* of the testimony cannot sensibly be denied. The hypotheses in  $\mathcal{K}_h$  are just that, and stories or arguments based on hypotheses can be undermined by other stories or arguments (see definition 6).

The combination of a knowledge base and inference rules allows us to build arguments and stories, derivations in the argumentation system. The integrated theory treats stories as *causal arguments*, which will allow us to attack and compare them as if they were arguments (definition 6 and 8). In definition 5, for a given argument the function **Prem** returns all its premises, **Conc** returns its conclusion, **Sub** returns all its sub-arguments, **TopRule** returns the last inference rule applied in the argument, **ERule** returns all its evidential arguments and **CRule** returns all its causal arguments.

**DEFINITION 5. [Arguments]** An *argument*  $A$  on the basis of a knowledge base  $KB$  in an argumentation system  $AS$  is:

1.  $\varphi$  if  $\varphi \in \mathcal{K}$ .  
 $\text{Prem}(A) = \{\varphi\}$ ;  $\text{Conc}(A) = \varphi$ ;  $\text{Sub}(A) = \{\varphi\}$ ;  $\text{TopRule}(A) = \text{undefined}$ ;  $\text{ERule}(A) = \emptyset$   $\text{CRule}(A) = \emptyset$
2.  $A_1, \dots, A_n \rightarrow \psi$  if  $A_1, \dots, A_n$  are arguments s.t. there exists a strict rule  $\text{Conc}(A_1), \dots, \text{Conc}(A_n) \rightarrow \psi$  in  $\mathcal{R}_s$ .

$\text{Prem}(A) = \text{Prem}(A_1) \cup \dots \cup \text{Prem}(A_n)$ ;  $\text{Conc}(A) = \psi$ ;  
 $\text{Sub}(A) = \text{Sub}(A_1) \cup \dots \cup \text{Sub}(A_n) \cup \{A\}$ ;  
 $\text{TopRule}(A) = \text{Conc}(A_1), \dots, \text{Conc}(A_n) \rightarrow \psi$ ;  
 $\text{ERule}(A) = \text{ERule}(A_1) \cup \dots \cup \text{ERule}(A_n)$ ;  
 $\text{CRule}(A) = \text{CRule}(A_1) \cup \dots \cup \text{CRule}(A_n)$

3.  $A_1, \dots, A_n \Rightarrow_C \psi$  if  $A_1, \dots, A_n$  are arguments s.t. there exists a causal rule  $\text{Conc}(A_1), \dots, \text{Conc}(A_n) \Rightarrow_C \psi$  in  $\mathcal{R}_c$ .  
 $\text{Prem}(A) = \text{Prem}(A_1) \cup \dots \cup \text{Prem}(A_n)$ ;  $\text{Conc}(A) = \psi$ ;  
 $\text{Sub}(A) = \text{Sub}(A_1) \cup \dots \cup \text{Sub}(A_n) \cup \{A\}$ ;  
 $\text{TopRule}(A) = \text{Conc}(A_1), \dots, \text{Conc}(A_n) \Rightarrow_C \psi$ ;  
 $\text{ERule}(A) = \text{ERule}(A_1) \cup \dots \cup \text{ERule}(A_n)$ ;  
 $\text{CRule}(A) = \text{CRule}(A_1) \cup \dots \cup \text{CRule}(A_n) \cup \{\text{Conc}(A_1), \dots, \text{Conc}(A_n) \Rightarrow_C \psi\}$
4.  $A_1, \dots, A_n \Rightarrow_E \psi$  if  $A_1, \dots, A_n$  are arguments s.t. there exists an evidential rule  $\text{Conc}(A_1), \dots, \text{Conc}(A_n) \Rightarrow_E \psi$  in  $\mathcal{R}_e$ .  
 $\text{Prem}(A) = \text{Prem}(A_1) \cup \dots \cup \text{Prem}(A_n)$ ;  $\text{Conc}(A) = \psi$ ;  
 $\text{Sub}(A) = \text{Sub}(A_1) \cup \dots \cup \text{Sub}(A_n) \cup \{A\}$ ;  
 $\text{TopRule}(A) = \text{Conc}(A_1), \dots, \text{Conc}(A_n) \Rightarrow_E \psi$ ;  
 $\text{ERule}(A) = \text{ERule}(A_1) \cup \dots \cup \text{ERule}(A_n) \cup \{\text{Conc}(A_1), \dots, \text{Conc}(A_n) \Rightarrow_E \psi\}$ ;  
 $\text{CRule}(A) = \text{CRule}(A_1) \cup \dots \cup \text{CRule}(A_n)$

An argument  $A$  is *causal* if  $\text{CRule}(A) \neq \emptyset$  and  $\text{ERule}(A) = \emptyset$ , *evidential* if  $\text{ERule}(A) \neq \emptyset$  and  $\text{CRule}(A) = \emptyset$ , and *mixed* if  $\text{CRule}(A) \neq \emptyset$  and  $\text{ERule}(A) \neq \emptyset$ .

The type of argument is indicated as follows: causal arguments are named  $C_i$ , evidential arguments are named  $E_i$  and mixed arguments are named  $M_i$ .

Like in the hybrid theory, we can build causal defeasible arguments (i.e. *stories*). If we take  $\mathcal{R}_c = \{rc_1, \dots, rc_{10}\}$  and assume that  $(\mathbf{R} \text{ and } \mathbf{F} \text{ argued}) \in \mathcal{K}_h$  we can construct the following causal arguments.

$C_{r1}$ : <b>R and F argued</b>	$C_{m1}$ : <b>R and F argued</b>
$C_{r2}$ : $C_{r1} \Rightarrow_c \mathbf{R} \text{ pulled } \mathbf{g}$	$C_{m2}$ : $C_{m1} \Rightarrow_c \mathbf{M} \text{ pulled } \mathbf{g}$
$C_{r3}$ : $C_{r2} \Rightarrow_c \mathbf{R} \text{ shot } \mathbf{F}$	$C_{m3}$ : $C_{m2} \Rightarrow_c \mathbf{R} \text{ pushed } \mathbf{M}$
$C_{r4}$ : $C_{r3} \Rightarrow_c \mathbf{F} \text{ hit}$	$C_{m4}$ : $C_{m3} \Rightarrow_c \mathbf{M} \text{ shot } \mathbf{F}$
$C_r$ : $C_{r4} \Rightarrow_c \mathbf{F} \text{ died}$	$C_{m5}$ : $C_{m4} \Rightarrow_c \mathbf{F} \text{ hit}$
$C_{r5}$ : $C_{r3} \Rightarrow_c \text{casings}$	$C_m$ : $C_{m5} \Rightarrow_c \mathbf{F} \text{ died}$
	$C_{m6}$ : $C_{m4} \Rightarrow_c \neg \text{casings}$

Arguments  $C_r$  and  $C_m$  are the two stories  $S_1$  and  $S_2$  that were also rendered in figure 1, and causal arguments  $C_{r5}$  and  $C_{m6}$  are expanded versions of the two stories  $S_1'$  and  $S_2''$  presented in figure 6.

Because it is assumed that there are no causal cycles, there is no causal argument  $C$  where  $\text{Prem}(C) = \text{Conc}(C)$ . One difference between stories in the integrated theory and stories in the hybrid theory is that in the latter, a story can have more than one story consequence so, for example, both **F died** and **casings** could both be part of one story. Because the integrated theory treats stories as arguments, it is not possible to have more than one conclusion for a causal argument. However, overlapping arguments with different conclusions are considered as-a-whole in the integrated theory, namely through the idea of positions, which will be further discussed below (definition 9).

Evidential arguments based can also be constructed in the integrated theory. For example, if we take  $\mathcal{R}_e = \{re_1, \dots, re_7\}$  and assume **Coroner report**<sub>1</sub>  $\in \mathcal{K}_e$  we can construct the following evidential arguments.

$E_{c0}$ : **Coroner report**<sub>1</sub>  
 $E_{c1}$ :  $E_{c0} \Rightarrow_e \mathbf{F\ hit}$   
 $E_1$ :  $E_{c1} \Rightarrow_e \mathbf{R\ shot\ F}$      $E_2$ :  $E_{c1} \Rightarrow_e \mathbf{M\ shot\ F}$

Because there are no causal cycles and the evidential rules are all based on causal relations, there is no evidential argument  $E$  where  $\text{Prem}(E) = \text{Conc}(E)$ .

Notice that evidential and causal arguments can be based on both evidence or hypotheses (i.e. any element of  $\mathcal{K}$  can serve as the premise of an any argument), whereas in the hybrid theory evidential arguments can only be based on evidence and causal stories can only be based on hypotheses. This is one example of how stories and arguments are further integrated. Another example is that the integrated theory allows for mixed causal-evidential (C-E) arguments. Assume that  $\mathcal{R}_c$  and  $\mathcal{R}_e$  are as above, and that **Women testimony**<sub>3</sub>  $\in \mathcal{K}_e$ .

$E_{w0}$ : **Women testimony**<sub>3</sub>     $M_1$ :  $E_{w1} \Rightarrow_c \mathbf{casings}$   
 $E_{w1}$ :  $E_{w0} \Rightarrow_e \mathbf{R\ shot\ F}$      $M_2$ :  $E_{w1} \Rightarrow_c \mathbf{F\ hit}$   
 $M_3$ :  $M_2 \Rightarrow_e \mathbf{M\ shot\ F}$

Definition 5 allows for all types of C-E reasoning:  $M_1$  (figure 3) is a correct way of reasoning with causal and evidential information. The reasoning captured by the mixed argument  $M_3$  (figure 4), however, is a classic example of C-E reasoning that is typically incorrect: as Pearl notes [12], an evidential rule  $e \Rightarrow_e c_2$  should only be invoked if no other explanation  $c_1$  of  $e$  is believed. In  $M_3$ , however, we already have explanation **R shot F** for **F hit**, so we should not be able to use  $re_7$  to subsequently infer **M shot F**. In other words, the application of rule  $re_7$  should be suppressed by the earlier application of causal rule  $rc_3$  in  $M_2$ .

There are essentially two ways to capture the suppression of evidential rules. One way is to specifically disallow an evidential inference rule to be applied to a conclusion that has been causally inferred; this is the way Pearl proposes in his C-E system [12]. However, in the integrated theory a cause that has been evidentially inferred (in the example: **M shot F**) is considered as a *competing alternative* to any other cause (e.g., **R shot F**). One reason to opt for this second way of modelling is that, as Pearl also notes, not all causal rules may present a clear separate cause that completely suppresses the inference of alternatives. That is, we may still want to infer the possible alternatives and then decide whether or not the original explanation suppresses any other alternatives.

### 3.4 Attack

The definition of alternatives is important for an integrated theory of stories and arguments, because in explanatory, abductive-causal reasoning this is where the dialectical component lies: two explanations (whether causal stories or evidential arguments) are in competition of they represent alternative causes for the same consequence. Hence, it can be said that alternative arguments *attack* each other.

In  $\text{ASPIC}^+$  arguments can be attacked in three ways: on their premises (undermining attack), on their conclusion (rebutting attack) or on an inference step (undercutting attack). The latter two are only possible on applications of defeasible inference rules. The integrated theory now adds another type of attack, namely *alternative attack*.

**DEFINITION 6. [Attack]**  $A$  attacks  $B$  iff  $A$  *undercuts*, *rebutts*, *undermines* or *alternative-attacks*  $B$ , where:

- $A$  *undercuts* argument  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = -n(r)$  and  $B' \in \text{Sub}(B)$  such that  $B'$ 's top rule  $r$  is defeasible.
- $A$  *rebutts* argument  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = -\varphi$  for some  $B' \in \text{Sub}(B)$  of the form  $B'' \Rightarrow_{C/E} \varphi$ .
- Argument  $A$  *undermines*  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = -\varphi$  for some  $B' = \varphi$ ,  $\varphi \notin \mathcal{K}_e$ .
- Argument  $A$  *alternative-attacks*  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = \varphi$  for some  $B' \in \text{Sub}(B)$  with  $\text{Conc}(B') = \psi$  and  $\varphi$  and  $\psi$  are alternatives according to definition 3.

Because definition 6 does not specifically distinguish causal and evidential arguments, we can have any combination of arguments or stories attacking each other in the different ways specified in definition 6. So, for example, the arguments  $C_{r5}$  and  $C_{m6}$  attack each other because they have opposite conclusions **casings** and **-casings**. Alternatives also attack each other, and because according to definition 3 they can be defined between any combination of evidential and causal rules, it is also possible to have, for example, two alternative evidential arguments that attack each other such as  $E_1$  and  $E_2$  (cf. figure 5). Further alternatives are, for example, the argument  $C_{r3}$  that **R shot F** attacks all the causal, evidential and mixed arguments in which **M shot F** (e.g.  $C_m$ ,  $E_2$ ,  $M_3$ ). Because the alternatives relation is symmetrical, we also have that, for example, the argument  $E_{c1}$  that **M shot F** attacks all arguments in which **R shot F**. Interestingly, argument  $M_3$  attacks its own subarguments  $E_{w1}$ ,  $M_2$  and is thus a self-attacking argument.

The above definition of attack puts evidential arguments and causal arguments (stories) on the same level: in the hybrid theory it was only possible for an evidential argument to attack either another argument or a causal story, but stories could not attack arguments or each other. Furthermore, evidential arguments that propose some causal relation now compete with causal stories that propose an alternative cause. Note that the important difference between arguments based on evidence and arguments based on hypotheses remains: the former cannot be undermined because the evidence itself cannot be attacked.

### 3.5 Argumentation semantics

Given an argumentation system (definition 1) plus a knowledge base (definition 4), we can now form an *integrated theory*, which induces an *integrated argumentation framework*.

**DEFINITION 7. [Integrated Argumentation Frameworks]** Let  $IT$  be an *integrated theory*  $(AS, KB)$ . An *integrated argumentation framework* (IAF) defined by  $AT$ , is a tuple  $\langle \mathcal{A}, \mathcal{C} \rangle$  where  $\mathcal{A}$  is the set of all finite arguments constructed from  $KB$  in  $AS$  and  $(X, Y) \in \mathcal{C}$  iff  $X$  attacks  $Y$ .

An integrated argumentation framework thus contains all the arguments and stories in a case and, most importantly, correctly indicates the attack relations between the contradictory and alternative arguments. Given such an IAF, we can define admissible sets and extensions as in [8] and determine whether an argument is justified [11].

**DEFINITION 8. [Argumentation Semantics]** Let  $\langle \mathcal{A}, \mathcal{C} \rangle$  be an IAF.

- A set of arguments  $S \subseteq \mathcal{A}$  is *conflict free* if there are no  $X, Y$  in  $S$  such that  $(X, Y) \in \mathcal{C}$ .

- An argument  $X \in \mathcal{A}$ ,  $X$  is *acceptable* w.r.t. some set of arguments  $S \subseteq \mathcal{A}$  iff for all arguments  $Y$  s.t.  $(Y, X) \in \mathcal{C}$  there is an argument  $Z \in S$  s.t.  $(Z, Y) \in \mathcal{C}$ .
- A conflict-free set of arguments  $S$  is *admissible* iff each argument  $X \in S$  is acceptable w.r.t.  $S$ .
- An admissible set of arguments  $S$  is a *complete* extension iff  $X \in S$  whenever  $X$  is acceptable w.r.t.  $S$ ;  $S$  is a *preferred* extension iff it is a set inclusion maximal complete extension;  $S$  is the *grounded* extension iff it is the set inclusion minimal complete extension; and  $S$  is a *stable* extension iff it is preferred and  $\forall Y \notin S, \exists X \in S$  s.t.  $(X, Y) \in \mathcal{C}$ .

For  $T \in \{\text{complete, preferred, grounded, stable}\}$ , an argument  $X$  is *sceptically* or *credulously* justified under the  $T$ -semantics if  $X$  belongs to all, respectively at least one,  $T$ -extension.

With the argumentation semantics, we can compare causal and evidential arguments in a case. As an example, let's go back to the main Rijkbloem case in figure 1. Assume that  $\mathcal{K}_h = \{\mathbf{R}$  and  $\mathbf{F}$  argued $\}$ ,  $\mathcal{K}_e$  contains the women's three testimonies ( $\mathbf{Women\ testimony}_{1,\dots,3}$ ), Rijkbloem's four testimonies ( $\mathbf{Rijkbloem\ testimony}_{1,\dots,4}$ ) and the coroner's three reports ( $\mathbf{Coroner\ report}_{1,\dots,3}$ ). Furthermore, assume that  $\mathcal{R}_c$  is as on page 4 and that  $\mathcal{R}_e$  contains all the rules on page 4 plus all the evidential inference rules for inferring conclusions from the testimonies and the reports ( $re_1, \dots, re_4$  are four examples of such rules). From this, we can construct the two main stories in the case  $C_r$  and  $C_m$  and their subarguments, as well as  $C_{r5}$  and  $C_{m6}$  (and subarguments) representing the possible predictions about the bullet casings at the scene of the crime (cf. figure 6). Furthermore, we have the evidential and mixed arguments based on the witness testimonies. There are then two preferred extensions, one that claims that father died because Rijkbloem shot him (i.e. containing all arguments except the ones using rule  $rc_3$  and the ones with *-casings* as a conclusion) and one that claims that father died because mother shot father (i.e. containing all arguments except the ones using  $rc_7$  and the ones with *casings* as a conclusion).

Whilst the alternative stories  $C_r$  and  $C_m$  and the contradictory stories  $C_{r5}$  and  $C_{m6}$  are (correctly) not in the same extension, the various sub-stories and the evidential arguments supporting the different stories are in both extensions because they are, considered outside of the context of their overarching alternative stories, not explicitly contradictory. So, for example, an extension which includes the story  $C_{r3}$ , that Rijkbloem shot Father, will not contain any stories or arguments arguing that it was mother who shot father. However, any sub-stories - for example, the story that mother pulled a gun - or evidential arguments - for example, the argument consisting of just  $\mathbf{women\ testimony}_3$  - will be in the same extension as  $C_{r3}$ . Thus, individual extensions of an IAF allow for fairly broad views on a case, capturing every element of a case that is not explicitly contradicted by the current assumptions.

### 3.6 Causal-Evidential constraints

Given these argumentation semantics, it is possible to show that the integrated theory conform to Pearl's C-E constraint [12] that it should in general not be possible to evidentially infer a conclusion from something that has itself been inferred using causal reasoning. More precisely:

**PROPOSITION 1. [C-E constraint]** Given an integrated argumentation framework  $\langle \mathcal{A}, \mathcal{C} \rangle$ , an argument  $A \in \mathcal{A}$  will not be (sceptically or credulously) justified under any semantics if there is an argument  $B \in \text{Sub}(A)$  such that  $B : B' \Rightarrow_e \chi$  and  $B' : B'' \Rightarrow_c \psi$  and  $\text{Conc}(B'') = \varphi$  where  $\varphi \neq \chi$ .

According to the definition of alternatives (definition 3),  $\varphi$  and  $\chi$  will be alternatives, which means that  $B''$  and  $B$  will attack each other (definition 6). Because  $B''$  is a subargument of  $B$ , this means that that argument  $B$  (and therefore also  $A$ ) will attack itself. Clearly, self-attacking arguments are not justified, since any set containing such arguments is not conflict-free.

In our example,  $M_3$  was an example of incorrect C-E reasoning. Given the above definitions of an IAF and the corresponding semantics,  $M_3$  self-attacks since it alternatively attacks its subargument  $E_{w1}$ . If we now consider an IAF where  $\mathcal{A} = \{E_{w0}, E_{w1}, M_2, M_3\}$  (page 4), we have the complete extension  $\{E_{w0}\}$ , which is also the grounded extension, and complete extension  $\{E_{w0}, E_{w1}, M_2\}$ , which is also the only preferred extension and the stable extension.

### 3.7 Positions

In the hybrid theory, the view on a case is often more narrow: the idea that an explanatory causal story needs to be subset-minimal only allows us to accept the directly relevant stories and arguments. While in some ways this does not provide the broad view that the extensions of the integrated theory do, it does provide more focus around the separate main stories. Hence, it would be informative to define *positions* in the integrated theory, admissible sets that present a meaningful and coherent view on the evidence and the facts of the case. For example, it makes sense that not all the arguments based on the women's testimonies and the arguments based on Rijkbloem's testimonies are in the same position, as, even though they do not directly contradict each other, they support different (alternative) explanations.

The way to look at positions is to consider a particular preferred extension in a case and determine what makes it different from other preferred extensions. In other words, what is the main focus of this view on the case?

**DEFINITION 9. [Positions]** Given an integrated argumentation framework IAT with one preferred extension  $PE$ , the *position*  $P$  is equal to  $PE$ . For an IAT with multiple preferred extensions  $PE_1, \dots, PE_n$ , each *position* is a set  $P_i \subseteq PE_i$  such that for each argument  $A \in P_i$ .

- $\exists B$  s.t.  $A \in \text{Sub}(B)$  and  $B \in PE_i$  and  $B \notin PE_j$  ( $i \neq j$ ); or
- $\neg \exists B$  s.t.  $A \notin \text{Sub}(B')$  and  $B' \notin PE_i$  and  $B' \in PE_j$  ( $i \neq j$ ).

Definition 9 effectively ensures that all the arguments in a position are either unique to the associated preferred extension or that they are not part of an argument that is unique to another extension. A position does not include the subarguments of arguments that are contradictory to the current position and thus lays the focus the specific arguments that make the associated extension unique. So, for example, any position which contains the main story that Rijkbloem shot father ( $C_r$ ) will contain the arguments that support this view but not the arguments that are part of the



opposite view, that mother shot father (e.g. the argument  $C_{m2}$  that mother pulled a gun).

### 3.8 Comparing positions

Because alternative-attack is symmetrical, we will often end up with different explanations of the evidence that are incompatible, different positions. The question is then how to compare these positions. Here it might be interesting to recall the criteria and concepts used to compare the stories in the original hybrid theory, like evidential support or evidential contradiction: which evidence from  $\mathcal{K}_e$  directly supports or contradicts an element in a position? The evidential support for a position would then be, for example, the set of evidence that are the premise of an argument in that position. Like in the hybrid theory, we can then use these concepts to compare positions and thus indicate preferences between positions; for example, the position with the larger (set-inclusive) evidential support is preferred. In this way, we can define preferences not only between arguments but also between particular *sets of arguments* [1].

## 4. DISCUSSION AND RELATED WORK

Aside from the hybrid theory [4], there have been other attempts to combine causal and evidential reasoning, most notably by Poole [15, 16]. The hybrid and integrated theories are based on these early ideas, but they expand on them significantly. For example, because these frameworks were originally intended for automatic diagnosis within relatively small and pre-defined domains, they do not incorporate aspects like reasoning about the rules that comprise the scenarios and more fine-grained scenario comparison criteria, and hence they are less suited to modelling and supporting more complex tasks such as analysing and making sense of the evidence in large criminal cases [17].

Interestingly, Poole’s abductive-causal approach has also been modelled as an argumentation framework in [6]. The interplay between stories and arguments and causal and evidential rules is, however, not directly captured by these frameworks. Another relevant argumentation-theoretic approach - though one that stays exclusively at the level of abstract argumentation - is proposed by [19], who combine explanatory reasoning with argumentation. An interesting avenue for future research would be to investigate how their formal account of the explanatory power of arguments relates to the arguments and positions in the integrated theory.

One approach which does capture the different subtle influences in causal and evidential reasoning is the Bayesian approach [13, 22, 20]. In this approach, which has steadily gained influence since the mid 1990’s at the expense of logical models such as [16, 12], the stories and arguments in a case are modelled as a Bayesian Network, a graphical representation of a joint probability distribution, where the nodes are variables (e.g., events, evidence) and the links represent the (in)dependencies between the variables. The directed links in a Bayesian network can be interpreted as causal rules, though a less strict criterion of ‘correlation’ is often also used.

Because the integrated theory is a logical interpretation of reasoning with causal information, it necessarily abstracts from many subtleties of causal inference captured by Bayesian reasoning. For example, the integrated framework does not capture the idea of accruing causes, or ‘explaining in’ as it

is sometimes called, where two causes  $c_1$  and  $c_2$  individually explain a consequence  $e$  but taken together make the probability of  $e$  even higher.

A valid question is then: why pursue the logical approach at all? Are we not merely building a, as Pearl [12] notes, ‘coarse logical abstraction’ of the type of reasoning that is easily captured by Bayesian inference? One answer is that one of the goals of the integrated theory is to capture and facilitate the way in which investigators and decision makers reason in real cases. Much of our reasoning is qualitative, and it therefore makes sense to capture it in a qualitative framework. In fact, one major shortcomings of Bayesian reasoning is that, in order to calculate the probability of the claims represented in the network, many conditional and prior probabilities are needed, which are often simply not available. Furthermore, these probabilities and the relations between them make Bayesian approaches difficult to understand for people that are not mathematically schooled, like police investigators and judges.

Bayesian approaches often struggle with the same issues regarding knowledge representation as logical approaches do: how do we represent stories and arguments, and how do we capture the interplay between them? Interestingly, many of the tentative proposed solutions to these questions originate from qualitative, (semi-)formal logical approaches. For example, the structures for Bayesian arguments proposed by [9] draw from work in argumentation theory [24], and the ‘scenario schemes’ for Bayesian networks in [22] were directly preceded by logical approaches to story schemes based on early qualitative work in AI [18]. Hence, it seems that a formal, logical theory is a good intermediary between on the one hand the informal, more natural conceptions of stories and arguments and on the other hand the quantitative mathematical models. This cross-fertilization between probabilistic and logical approaches is also the subject of recent work in argumentation [21, 20].

## 5. CONCLUSIONS AND FUTURE RESEARCH

This paper has expanded and further formalised the integrated theory for reasoning with evidence that was first proposed in a recent short paper [2]. In this integrated theory, it is possible to reason with both causal and evidential information, and to employ stories (causal arguments) and evidential arguments in one’s reasoning whenever needed. Thus, it offers the ‘best of both worlds’ with respect to story-based reasoning [23, 14] and argumentative reasoning [3] in the process of proof.

The integrated theory offers a dialectical framework for abductive argumentation which instantiates the *ASPIC*<sup>+</sup> framework for structured argumentation. In addition to defining a combination of abductive and non-abductive argumentation, the integrated theory also correctly captures argumentative causal reasoning. Furthermore, the positions (definition 9) represent coherent views on a case which can be compared using different criteria. Thus, this adds a new way of defining preferences between sets of arguments in an argumentation framework.

The integrated theory is a natural continuation of the work on the hybrid theory [4], which was the first formal theory that combined causal stories and evidential arguments. However, as the name suggests, the hybrid theory stops short of truly integrating stories and arguments into a single theory. The integrated theory proposed in the cur-

rent paper therefore improves on the hybrid theory in a few important ways.

- It is now possible to construct *mixed* causal-evidential arguments as well as causal stories and evidential arguments. This mixed causal-evidential reasoning adheres to Pearl's [12] constraints for C-E reasoning.
- In addition to alternative causal stories, which were part of the hybrid theory, the integrated theory also defines alternative evidential and mixed arguments. Thus, it is possible to emphasize causal or evidential parts of a case as needed whilst keeping the case outcome the same.
- The integrated theory treats stories as causal arguments that can attack not only alternative explanations of the evidence but also contradictory stories (e.g. alibi stories). Thus arguments and stories can both be compared using the common semantics for argumentation [8].

As for future research, one of the most urgent and interesting subjects is perhaps the exploration of how various other phenomena that play an important role in causal reasoning can be captured in a logical framework, such as accrual and the distinction between 'strong' causal rules, which preclude any subsequent evidential inference (as in proposition 1) and 'weak' causal rules, which allow evidential inferences to be made after they have been applied. Such an extended logical framework can then serve as a gateway between natural, informal reasoning with causal stories and evidential arguments and more probabilistic and mathematically inclined accounts of causal reasoning such as Bayesian networks.

## 6. REFERENCES

- [1] L. Amgoud and S. Vesic. Rich preference-based argumentation frameworks. *International Journal of Approximate Reasoning*, 55(2):585–606, 2014.
- [2] F. Bex. Towards an integrated theory of causal scenarios and evidential arguments. In *Proceedings of COMMA 2014, Frontiers in Artificial Intelligence and Applications*, volume 266, pages 11–18.
- [3] F. Bex, H. Prakken, C. Reed, and D. Walton. Towards a Formal Account of Reasoning about Evidence: Argumentation Schemes and Generalisations. *Artificial Intelligence and Law*, 11(2/3):125–165, 2003.
- [4] F. Bex, P. van Koppen, H. Prakken, and B. Verheij. A hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law*, 2:123–152, 2010.
- [5] F. Bex and D. Walton. Combining explanation and argumentation in dialogue. In *Selected papers from the 2011-2013 CMNA Workshops, Lecture Notes in Artificial Intelligence*, 2012. to appear.
- [6] A. Bondarenko, P. Dung, R. Kowalski, and F. Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93:63–101, 1997.
- [7] S. v. d. Braak, H. v. Oostendorp, H. Prakken, and G. Vreeswijk. Representing narrative and testimonial knowledge in sense-making software for crime analysis. In *Proceedings of JURIX 2008*, pages 160–169. IOS Press, 2008.
- [8] P. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and  $n$ -person games. *Artificial Intelligence*, 77:321–357, 1995.
- [9] D. A. Lagnado, N. E. Fenton, and M. Neil. Legal idioms: a framework for evidential reasoning. *Argument and Computation*, pages 1–18, 2012.
- [10] P. Lucas. Symbolic diagnosis and its formalisation. *The Knowledge Engineering Review*, 12(02):109–146, 1997.
- [11] S. Modgil and H. Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397, 2013.
- [12] J. Pearl. Embracing causality in default reasoning. *Artificial Intelligence*, 35:259–271, 1988.
- [13] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.
- [14] N. Pennington and R. Hastie. Reasoning in explanation-based decision making. *Cognition*, 49(1–2):123–163, 1993.
- [15] D. Poole. Explanation and prediction: an architecture for default and abductive reasoning. *Computational Intelligence*, 5(2):97–110, 1989.
- [16] D. Poole. Representing diagnosis knowledge. *Annals of Mathematics and Artificial Intelligence*, 11(1-4):33–50, 1994.
- [17] H. Prakken and S. Renooij. Reconstructing causal reasoning about evidence: a case study. *Legal knowledge and information systems. JURIX*, pages 131–142, 2001.
- [18] R. Schank and R. Abelson. *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*. Lawrence Erlbaum, Hillsdale, NJ, 1977.
- [19] D. Šešelja and C. Straßer. Abstract argumentation and explanation applied to scientific debates. *Synthese*, 190(12):2195–2217, 2013.
- [20] S. Timmer, J.-J. C. Meyer, H. Prakken, S. Renooij, and B. Verheij. Extracting legal arguments from forensic bayesian networks. In *Proceedings of the JURIX 2014 Conference*, pages 71–80. IOS Press, 2014.
- [21] B. Verheij. Arguments and their strength: Revisiting Pollock's anti-probabilistic starting points. In *Computational Models of Argument. Proceedings of COMMA 2014*, pages 433–444. IOS Press, 2014.
- [22] C. Vlek, H. Prakken, S. Renooij, and B. Verheij. Extracting scenarios from a Bayesian Network as explanations for legal evidence. In *Proceedings of the JURIX 2014 conference*, pages 150–159. IOS Press, New York (New York), 2014.
- [23] W. Wagenaar, P. Van Koppen, and H. Crombag. *Anchored Narratives. The Psychology of Criminal Evidence*. Harvester Wheatsheaf, London, 1993.
- [24] D. Walton. *Argumentation Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, 1996.